

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 18:45:47 ; Search time 2.90323 Seconds
(without alignments)
128.577 Million cell updates/sec

Title: US-09-674-593-3
Perfect score: 57
Sequence: 1 LPRWPPQL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	44	77.2	1733	1	VNUA_PRVKA	P33485 pseudorabies
2	41	71.9	410	1	KYK2_DICDI	P18161 dictyosteli
3	41	71.9	789	1	K6PF_HAECO	Q27665 haemochus
4	40	70.2	1274	1	ENAM_MOUSE	O55196 mus musculus
5	39	68.4	197	1	Y089_MYCTU	Q10886 mycobacteri
6	39	68.4	280	1	CC2B_ANTNA	Q38773 antirrhinum
7	39	68.4	715	1	S141_HUMAN	Q92503 homo sapien
8	39	68.4	852	1	WS14_HUMAN	Q9np71 homo sapien
9	38	66.7	184	1	T13C_HUMAN	Q96rj3 homo sapien
10	38	66.7	310	1	Y497_MYCTU	Q11162 mycobacteri
11	38	66.7	720	1	AB13_ARATH	Q01593 arabidopsis
12	38	66.7	1242	1	WDRA_HUMAN	Q9hb66 homo sapien
13	38	66.7	2411	1	DAB_DROME	P98081 drosophila
14	37.5	65.8	5035	1	RYR1_PIG	P16960 sus scrofa
15	37.5	65.8	5037	1	RYR1_RABIT	P11716 oryctolagus
16	37	64.9	13	1	TP13_PHYRO	P40096 phyllomedusa
17	37	64.9	194	1	HLV2_ACTPL	P46028 actinobacil
18	37	64.9	265	1	COX3_MYTEL	P41775 mytilus edu
19	37	64.9	294	1	CC2_VIGUN	P52389 vigna ungui
20	37	64.9	340	1	MPCP_CAEEL	P40614 caenorhabdi
21	37	64.9	356	1	MPCP_RAT	P16036 rattus norv
22	37	64.9	357	1	MPCP_MOUSE	Q8vem8 mus musculu
23	37	64.9	362	1	MPCP_BOVIN	P12234 bos taurus
24	37	64.9	362	1	MPCP_HUMAN	Q00325 homo sapien
25	37	64.9	369	1	CYRG_MOUSE	P34902 mus musculu
26	37	64.9	405	1	FUT4_HUMAN	P22083 homo sapien
27	37	64.9	433	1	FUT4_RAT	Q62994 rattus norv
28	37	64.9	485	1	SOC7_HUMAN	O14512 homo sapien
29	37	64.9	486	1	GATB_AERPE	Q9yb79 aeropyrum p
30	37	64.9	509	1	RA18_MOUSE	Q9qkx2 mus musculu
31	37	64.9	515	1	ENV_BLV	P51519 bovine leuk
32	37	64.9	515	1	ENV_BLVAF	P25504 bovine leuk
33	37	64.9	515	1	ENV_BLVAV	P25057 bovine leuk

34	37	64.9	515	1	ENV_BLVB2	P25506 bovine leuk
35	37	64.9	515	1	ENV_BLVBS	P25507 bovine leuk
36	37	64.9	515	1	ENV_BLVJ	P03380 bovine leuk
37	37	64.9	536	1	C7C2_MAIZE	Q43255 zea mays (m
38	37	64.9	614	1	ACES_RAT	P37136 rattus norv
39	37	64.9	767	1	ACES_CHICK	P36196 gallus gall
40	37	64.9	1001	1	PTPX_MOUSE	P80560 mus musculu
41	37	64.9	1130	1	C2TA_HUMAN	P33076 homo sapien
42	37	64.9	1173	1	GIGA_ORYSA	Q9awl7 oryza sativ
43	37	64.9	1234	1	PIP3_MOUSE	P51432 mus musculu
44	37	64.9	1395	1	CUTL_MOUSE	P53564 mus musculu
45	37	64.9	1783	1	RAA3_CHLRE	O9fec4 chlamydomon

ALIGNMENTS

RESULT 1
VNUA_PRVKA
ID VNUA_PRVKA STANDARD; PRT; 1733 AA.
AC P33485;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Probable nuclear antigen.
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=33703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021039; PubMed=2171211;
RA Vleck C., Kozmik Z., Paces V., Schirm S., Schwyzler M.;
RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
oriented open reading frame: characterization of their promoter and
enhancer regions.";
RL Virology 179:365-377(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34651; AAA47471.1; -;
DR PIR; B45344; B45344.
FT DOMAIN 112 117 POLY-THR.
FT DOMAIN 179 1733 GLY-RICH.
FT DOMAIN 192 196 POLY-SER.
FT DOMAIN 271 298 POLY-PRO.
FT DOMAIN 304 308 POLY-ARG.
FT DOMAIN 883 889 POLY-GLY.
FT DOMAIN 1398 1405 POLY-GLY.
SQ SEQUENCE 1733 AA; 172166 MW; 0C8CD8BE475BB5E2 CRC64;

Query Match 77.2%; Score 44; DB 1; Length 1733;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PRWPPP 7
DB 213 PRWPPP 218

RESULT 2
KYK2_DICDI
ID KYK2_DICDI STANDARD; PRT; 410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYKB OR DPYK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33784; AAA33203.1; -
CC PIR; B35670; B35670.
CC HSP; P08631; 1AD5.
CC DictyDb; D003011; pyKB.
CC InterPro; IPR000719; Euk_pkinase.
CC TrEMBL; P8004040; STY_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AEC1 CRC64;

Query Match 71.9%; Score 41; DB 1; Length 410;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPQL 9
Db 273 PRWRPPEL 280
||| ||:|

RESULT 3
K6PF_HABCO STANDARD; PRT; 789 AA.
AC Q27665;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (Phosphohexokinase).
GN PFK.
OS Haemochus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemochidae; Haemochinae; Haemochus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92140452; PubMed=1838137;
RA Klein R.D., Faureau M.A., Winterrowd C.A., Hatzenbuehler N.T.,
RA Shea M.H., Nulf S.C., Geary T.G., Olson E.R.;

Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYKB OR DPYK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33784; AAA33203.1; -
CC PIR; B35670; B35670.
CC HSP; P08631; 1AD5.
CC DictyDb; D003011; pyKB.
CC InterPro; IPR000719; Euk_pkinase.
CC TrEMBL; P8004040; STY_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AEC1 CRC64;

Query Match 71.9%; Score 41; DB 1; Length 410;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPQL 9
Db 273 PRWRPPEL 280
||| ||:|

RESULT 3
K6PF_HABCO STANDARD; PRT; 789 AA.
AC Q27665;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
DE (Phosphohexokinase).
GN PFK.
OS Haemochus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemochidae; Haemochinae; Haemochus.
OX NCBI_TaxID=6289;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92140452; PubMed=1838137;
RA Klein R.D., Faureau M.A., Winterrowd C.A., Hatzenbuehler N.T.,
RA Shea M.H., Nulf S.C., Geary T.G., Olson E.R.;

"Cloning of a cDNA encoding phosphofructokinase from Haemochus
contortus."
Mol. Biochem. Parasitol. 48:17-26(1991).
-!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
fructose 1,6-bisphosphate.
-!- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
-!- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY. TWO DOMAIN
SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M59805; AAA29181.1; -
HSP; P06998; 2PFK.
InterPro; IPR000023; Ppfruckinase.
Pfam; PF00365; PFK; 2.
PRINTS; PR00476; PPFRTKINASE.
ProDom; PD000707; Ppfruckinase; 2.
PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 2.
KW Kinase; Transferase; Glycolysis; Repeat.
SQ SEQUENCE 789 AA; 87184 MW; 1245288AE2289BCF CRC64;

Query Match 71.9%; Score 41; DB 1; Length 789;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPP 7
Db 246 IPEWPPP 252
||| |||||

RESULT 4
ENAM_MOUSE STANDARD; PRT; 1274 AA.
ID O55196;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Enamelin precursor.
GN ENAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Swiss Webster; TISSUE=Enamel epithelium;
RX MEDLINE=20514993; PubMed=11062988;
RA Hu C.-C., Simmer J.P., Bartlett J.D., Nanci A., Qian Q., Zhang C.,
RA Ryu O.H., Xue J., Fukae M., Uchida T., McDougall M.;
RT "Murine enamel: cDNA and derived protein sequences."
RL Connect. Tissue Res. 39:47-61(1998).
CC -!- FUNCTION: PEPTIDES DERIVED FROM THE PARENT ENAMELIN ARE COMPONENTS
CC OF ENAMEL, A UNIQUE AND HIGHLY MINERALIZED ECTODERMAL TISSUE
CC COVERING VERTEBRATE TEETH (BY SIMILARITY).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U82698; AAB94312.1; -
MGD; MGI:133372; Enam.
Signal; Enamel; Glycoprotein.
KW Signal; Enamel; Glycoprotein.
FT SIGNAL 1 38

```

```
FT CHAIN 39 1274 ENAMELIN.
FT CARBOHYD 130 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1274 AA; 140954 MW; F9DBD1CC9D327143 CRC64;

Query Match 70.2%; Score 40; DB 1; Length 1274;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRPWPPP 7
DB 97 MPWPPPP 103

RESULT 5
Y089_MYCTU STANDARD; PRT; 197 AA.
AC Q10886;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative methyltransferase RV0089 (EC 2.1.1.-).
GE RV0089 OR MT0098 OR MTCY251.07.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98299987; PubMed=9634230;
RA Gordon S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Cole S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF00398; RnaAD; 1.
KW Hypothetical protein; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 197 AA; 21596 MW; 004334F1A289EC51 CRC64;

Query Match 68.4%; Score 39; DB 1; Length 197;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 RWPPPQ 8
DB 158 KWPPPQ 163

RESULT 6
CC2B_ANTMA STANDARD; PRT; 280 AA.
AC Q38773;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cell division control protein 2 homolog B (EC 2.7.1.1-) (Fragment).
GS CDC2B.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Flower;
RX MEDLINE=96434535; PubMed=8837502;
RA Robert P.R., Gaudin V., Lunness P., Coen E.S., Doonan J.H.;
RT "Distinct classes of cdc2-related genes are differentially expressed
RT during the cell division cycle in plants.";
RL Plant Cell 8:1465-1476(1996).
CC -!- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT TYR-2 INACTIVATES THE
CC ENZYME, WHILE PHOSPHORYLATION AT THR-148 ACTIVATES IT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDKX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match 68.4%; Score 39; DB 1; Length 280;
Best Local Similarity 62.5%; Pred. No. 41;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPPOL 9
|:|:|:|:
Db 229 PKWPPREL 236

RESULT 7

S141_HUMAN STANDARD; PRT: 715 AA.
AC Q92503; Q99780;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SEC14-like protein 1.
GN SEC14L1 OR SEC14L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=96302338; PubMed=8697811;
RA Chinen K., Takahashi E., Nakamura Y.;
RT "Isolation and mapping of a human gene (SEC14L), partially homologous
RT to yeast SEC14, that contains a variable number of tandem repeats
RT (YNT) site in its 3' untranslated region.";
RL Cytogenet. Cell Genet. 73:218-223(1996).
RN [2]
RP SEQUENCE OF 423-715 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97264341; PubMed=9110174;
RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
RT "Large-scale concatenation cDNA sequencing.";
RL Genome Res. 7:353-358(1997).
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- SIMILARITY: CONTAINS 1 CRAL-TRIO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D67029; BAAL1048.1; -
CC EMBL: U79284; AAB50220.1; -
CC Genew; HGNC:10698; SEC14L1.
CC MIM; 601504; -
CC InterPro; IPR001251; CRAL-TRIO.
CC Pfam; PF00650; SEC14; 1.
CC SMART; SM00516; CRAL-TRIO; 1.
CC PROSITE; PS01091; CRAL-TRIO; 1.
CC DOMAIN 319 495 CRAL-TRIO.
CC CONFLICT 715 715 R -> RWRFC (IN REF. 2).
SQ SEQUENCE 715 AA; 81277 MW; F35B08CD8ADDE79A CRC64;

Query Match 68.4%; Score 39; DB 1; Length 715;
Best Local Similarity 55.8%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPRWPPPOL 9
:|:|:|:|:
Db 178 VPRWSPPSI 186

RESULT 8

WS14_HUMAN STANDARD; PRT: 852 AA.
ID Q9N7L1; Q9BY06; Q9BY05; Q9BY04; Q9BY03; Q96E48; Q9Y2P3;
AC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Williams-Beuren syndrome chromosome region 14 protein (WS basic-helix-
DE loop-helix leucine zipper protein) (WS-BHLH) (Mlx interactor).
GN WBSCR14 OR MIO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20241700; PubMed=10780788;
RA de Luis O., Valero M.C., Perez Jurado L.A.;
RT "WBSCR14, a putative transcription factor gene deleted in Williams-
RT Beuren syndrome: complete characterisation of the human gene and the
RT mouse ortholog.";
RL Eur. J. Hum. Genet. 8:215-222(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), AND SUBUNIT.
RX PubMed=11230181;
RA Cairo S., Merla G., Urbinati F., Ballabio A., Raymond A.;
RT "WBSCR14, a gene mapping to the Williams-Beuren syndrome deleted
RT region, is a new member of the Mlx transcription factor network.";
RL Hum. Mol. Genet. 10:617-627(2001).
RN [3]
RP SEQUENCE OF 620-852 FROM N.A. (ISOFORM 4).
RX MEDLINE=99075645; PubMed=9860302;
RA Meng X., Lu X., Li Z., Green E.D., Massa H., Trask B.J., Morris C.A.,
RA Keating M.T.;
RT "Complete physical map of the common deletion region in Williams
RT syndrome and identification and characterization of three novel
RT genes.";
RL Hum. Genet. 103:590-599(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Transcriptional repressor. Binds to the canonical and
CC non-canonical E box sequences 5'-CACGTG-3' (by similarity).
CC -1- SUBUNIT: Binds DNA as a heterodimer with TCFL4/MLX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 6 ISOFORMS; 1/ALPHA (SHOWN HERE), 2/BETA,
CC 3/GAMMA, 4/DELTA, 5/EPSILON AND 6; ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, HEART, KIDNEY, CEREBELLUM
CC AND INTESTINAL TISSUES.
CC -1- DISEASE: Haploinsufficiency of WBSCR14 may be the cause of certain
CC cardiovascular and musculo-skeletal abnormalities observed in
CC Williams-Beuren syndrome (WBS), a rare developmental disorder. It
CC is a contiguous gene deletion syndrome involving genes from
CC chromosome band 7q11.23.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. BHLH-ZIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF156673; AAF68176.1; -
CC EMBL: AF156603; AAF68174.1; -
CC EMBL: AF245470; AAK20935.1; -
CC EMBL: AF245471; AAK20936.1; -
CC EMBL: AF245472; AAK20937.1; -
CC EMBL: AF245473; AAK20938.1; -


```

DR EMBL; AF245474; AAK20939.1; -
DR EMBL; AF056184; AAD28084.1; -
DR EMBL; BC012925; AAH12925.1; -
DR HSSP; P25912; IHLO.
DR Genew; HGNC:12744; WBSR14.
DR MIN; 605678; -
DR MIN; 194050; -
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 4.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00888; HLH_2; 1.
KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
KW Williams-Beuren syndrome; Alternative splicing.
FT DOMAIN 387 394 POLY-PRO.
FT DOMAIN 409 417 POLY-PRO.
FT DNA_BIND 648 662 BASIC DOMAIN.
FT DOMAIN 688 702 HELIX-LOOP-HELIX MOTIF (POTENTIAL).
FT DOMAIN 703 724 LEUCINE-ZIPPER.
FT VARSPLIC 176 268 MISSING (IN ISOFORM 6).
FT VARSPLIC 558 575 QETVPEPCTFLPTTPAP -> AVNGCGQTSAPCQALGL
(IN ISOFORM 5).
FT VARSPLIC 576 852 MISSING (IN ISOFORM 5).
FT VARSPLIC 647 648 MISSING (IN ISOFORM 3 AND ISOFORM 4).
FT VARSPLIC 687 705 MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT VARSPLIC 814 815 TV -> ST (IN ISOFORM 6).
FT VARSPLIC 816 852 MISSING (IN ISOFORM 6).
FT CONFLICT 558 558 MISSING (IN REF. 4).
SQ SEQUENCE 852 AA; 93071 MW; D495EC3D7C0A72EC CRC64;

Query Match 68.4%; Score 39; DB 1; Length 852;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RWPPQ 8
Db 210 RWPPPE 215

RESULT 9
T13C_HUMAN
ID T13C_HUMAN STANDARD; PRT; 184 AA.
AC Q96R3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 13C (B cell-
activating factor receptor) (BAFF receptor) (BAFF-R) (Blys receptor
3).
DE TNFRSF13C OR BAFFR OR BR3.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
T13C-B-cell lymphoma;
RX MEDLINE=21442025; PubMed=11509692;
RA Thompson J.S., Bixler S.A., Qian F., Vora K., Scott M.L.,
Cachero T.G., Hession C., Schneider P., Sizing I.D., Mullen C.,
Strauch K., Zafari M., Benjamin C.D., Tschopp J., Browning J.L.,
Ambrose C.;
RA "BAFF-R, a newly identified TNF receptor that specifically interacts
with BAFF.";
RT Science 293:2108-2111(2001).
RN [2]
RN FUNCTION.
RX MEDLINE=21475520; PubMed=11591325;
RA Yan M., Brady J.R., Chan B., Lee W.P., Hsu B., Harless S.M.,
Cancro M.P., Grewal I.S., Dixit V.M.;
RA "Identification of a novel receptor for B lymphocyte stimulator that
is mutated in a mouse strain with severe B cell deficiency.";
RL Curr. Biol. 11:1547-1552(2001).

```

```

CC CC EMBL; AF373846; AAK91826.1; -
CC CC Promotes the survival of mature B-cells and the B-cell response.
CC CC - SUBCELLULAR LOCATION: Type III membrane protein (Probable).
CC CC - ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC CC produced by alternative splicing.
CC CC - TISSUE SPECIFICITY: Highly expressed in spleen and lymph node, and
CC CC in resting B-cells. Detected at lower levels in activated B-cells,
CC CC resting CD4+ T-cells, in thymus and peripheral blood leukocytes.
CC CC - SIMILARITY: CONTAINS 1 TNFR-CYS REPEAT.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; AF373846; AAK91826.1; -
CC CC Genew; HGNC:17755; TNFRSF13C.
CC CC MIN; 606269; -
CC CC PROSITE; PS00652; TNFR_NGFR_1; FALSE_NEG.
CC CC PROSITE; PS50050; TNFR_NGFR_2; FALSE_NEG.
KW Receptor; Immune response; Signal-anchor; Transmembrane;
KW Alternative splicing.
FT DOMAIN 1 78 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 79 99 SIGNAL-ANCHOR (TYPE III MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 100 184 CYTOPLASMIC (POTENTIAL).
FT REPEAT 18 35 TNFR-CYS (PARTIAL).
FT DISULFID 19 32 BY SIMILARITY.
FT DISULFID 24 35 BY SIMILARITY.
FT VARSPLIC 143 143 P -> PA (IN ISOFORM 2).
SQ SEQUENCE 184 AA; 18863 MW; F2BF98099A27138 CRC64;

Query Match 66.7%; Score 38; DB 1; Length 184;
Best Local Similarity 83.3%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RWPPPP 7
Db 143 PAWPPP 148

RESULT 10
Y497_MYCTU
ID Y497_MYCTU STANDARD; PRT; 310 AA.
AC Q11162;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv0497
GN RV0497 OR MT0517 OR MTCY20G9.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Hoilroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RT Nature 393:537-544(1998).
RL [2]
RN

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.LEPRAE ML2433.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 277162; CAB00923.1; -
CC EMBL: AE006952; AAK44740.1; -
CC TIGR: MT0517; -
CC TubercuList; RV0497; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT DOMAIN 33 39 POLY-ARG.
FT DOMAIN 197 202 POLY-ALA.
SQ SEQUENCE 310 AA; 33092 MW; 4954027F694DF5C2 CRC64;

Query Match 66.7%; Score 38; DB 1; Length 310;
Best Local Similarity 77.8%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 2 PRWP--PPQ 8
DB 113 PRWPKSPQP 121

RESULT 11
AB13_ARATH
ID AB13_ARATH STANDARD; PRT; 720 AA.
AC Q01593;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Absciscic acid-insensitive protein 3.
GN AB13 OR AT3G24650 OR MSD24.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=93076109; PubMed=1359917;
RA Giraudat J., Hauge B.M., Valon C., Smalle J., Parcy F.,
RA Goodman H.M.;
RT "Isolation of the Arabidopsis AB13 gene by positional cloning.";
RL Plant Cell 4:1251-1261(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221(2000).

```

```

CC -!- FUNCTION: COULD PARTICIPATE IN ABSICISIC ACID-REGULATED GENE
CC EXPRESSION DURING SEED DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: CONTAINS 1 TF-B3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X68141; CAA48241.1; -
CC EMBL: AB00740; BAB01214.1; -
CC PIR: J01676; J01676.
CC InterPro: IPR003340; TF_B3.
CC Pfam: PF02362; B3; 1.
CC Developmental protein; Transcription regulation; DNA-binding;
KW Activator; Nuclear protein.
FT DOMAIN 532 720 TF-B3.
SQ SEQUENCE 720 AA; 79500 MW; D6A0F740D733060F CRC64;

Query Match 66.7%; Score 38; DB 1; Length 720;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
DB 391 PSMWPPP 396

RESULT 12
WDRA_HUMAN
ID WDRA_HUMAN STANDARD; PRT; 1242 AA.
AC Q9HBG6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-repeat protein 10.
GN WDR10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21134653; PubMed=11242542;
RA Gross C., De Baere E., Lo A., Chang W., Messiaen L.;
RT "Cloning and characterization of human WDR10, a novel gene located at
RT 3q21 encoding a WD-repeat protein that is highly expressed in
RT pituitary and testis.";
RL DNA Cell Biol. 20:41-52(2001).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Expressed in many tissues. Predominant
CC expression in testis and pituitary.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF244930; AAG15427.1; -
CC Genew; HGNC:13556; WDR10.
CC MIM: 606045; -
CC InterPro: IPR001680; WD40.
CC Pfam: PF00400; WD40; 7.
CC ProDom; PD000018; WD40; 1.
CC PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.

```


CC structure interacts with the cytoplasmic region of the
CC dihydropyridine receptor.
CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-
CC release channel in junctional SR and modulates its activity.
CC -!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15209; CAA33279.1; -;
DR EMBL; X15750; CAA33762.1; -;
DR PIR; S04654; S04654;
DR InterPro; IPR000699; Ca-rel_channel.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000636; M-channel_nlg.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR003032; Ryr.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003878; SPRY_domain.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF00622; SPRY; 3.
DR Pfam; PF01365; RYDR_ITPR; 2.
DR Pfam; PF02026; Ryr; 4.
DR Pfam; PF02815; MIR; 4.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
KW Phosphorylation; Glycoprotein.
FT DOMAIN 1 3123
FT TRANSMEM 3124 3144
FT M' (POTENTIAL).
FT TRANSMEM 3188 3206
FT M'' (POTENTIAL).
FT TRANSMEM 3985 4004
FT M1 (POTENTIAL).
FT TRANSMEM 4023 4041
FT M2 (POTENTIAL).
FT TRANSMEM 4277 4300
FT M3 (POTENTIAL).
FT TRANSMEM 4342 4362
FT M4 (POTENTIAL).
FT TRANSMEM 4559 4580
FT M5 (POTENTIAL).
FT TRANSMEM 4648 4671
FT M6 (POTENTIAL).
FT TRANSMEM 4789 4809
FT M7 (POTENTIAL).
FT TRANSMEM 4837 4856
FT M8 (POTENTIAL).
FT TRANSMEM 4879 4898
FT M9 (POTENTIAL).
FT TRANSMEM 4914 4937
FT M10 (POTENTIAL).
FT SPRY 1.
FT DOMAIN 659 797
FT SPRY 2.
FT DOMAIN 1085 1208
FT SPRY 3.
FT DOMAIN 1430 1570
FT GLU-RICH (ACIDIC).
FT DOMAIN 1873 1913
FT REPEAT 842 2959
FT 6 X APPROXIMATE REPEATS.
FT REPEAT 842 955
FT 1.
FT REPEAT 956 1069
FT 2.
FT REPEAT 1345 1360
FT 3 (INCOMPLETE).
FT REPEAT 1373 1388
FT 4 (INCOMPLETE).
FT REPEAT 2726 2845
FT 5.
FT REPEAT 2846 2959
FT 6.
FT MOD_RES 2843 2843
FT PHOSPHORYLATION (BY PKA AND PKG).
FT MOD_RES 3952 3952
FT PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 4323 4323
FT WITH ACHR-SUBUNIT'S TRANSMEMBRANE SEGMENTS
FT SIMILAR 4628 4861
FT M2 AND M3.
FT CARBOHYD 3466
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3909
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3950
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4149
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4864
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 2015 2015
FT E -> D (IN REF. 2).
FT CONFLICT 3481 3485
FT MISSING (IN REF. 2).
SQ SEQUENCE 5037 AA; 565238 MW; 4ABD87AA26697070 CRC64;

Query Match 65.8%; Score 37.5; DB 1; Length 5037;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 7; Gaps 1;
Qy 1 LPRW-----PPQQL 9
Db 3281 LPRWWRGPEAPPAL 3296
Search completed: November 26, 2002, 12:02:26
Job time : 7.90323 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 11:58:11 ; Search time 4.93548 Seconds
(without alignments)
175.304 Million cell updates/sec

Title: US-09-674-593-3
Perfect score: 57
Sequence: 1 LPRWPPPPQL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	77.2	1733	1 B45344	probable nuclear a
2	44	77.2	1958	2 B40505	hypothetical prote
3	42	73.7	211	2 C84681	hypothetical prote
4	41	71.9	190	2 F75336	conserved hypoteth
5	41	71.9	211	2 T49803	hypothetical prote
6	41	71.9	410	2 B35670	protein-tyrosine k
7	41	71.9	789	1 A45617	6-phosphofructokin
8	40	70.2	61	2 T03587	isocitrate dehydro
9	40	70.2	166	2 E90819	hypothetical prote
10	40	70.2	180	2 T25460	hypothetical prote
11	40	70.2	206	2 A87253	hypothetical prote
12	40	70.2	218	2 G72754	probable phosphoen
13	40	70.2	409	2 T24543	hypothetical prote
14	40	70.2	578	2 T22888	hypothetical prote
15	40	70.2	592	2 T22885	hypothetical prote
16	40	70.2	629	2 T22890	hypothetical prote
17	40	70.2	1274	2 T37193	enamelin matrix pr
18	39	68.4	197	2 A70750	hypothetical prote
19	39	68.4	280	2 T17116	protein kinase cdc
20	39	68.4	299	2 AF3045	transcription regu
21	39	68.4	299	2 F98240	gstR protein (AF00
22	39	68.4	613	2 JC7762	SOX-3 protein - gu
23	39	68.4	669	2 T51246	ARL1 protein [limp
24	38	66.7	104	2 S53993	hypothetical prote
25	38	66.7	153	2 G72481	probable histidine
26	38	66.7	304	2 B69304	3-hydroxyacyl-CoA
27	38	66.7	310	2 D70745	hypothetical prote
28	38	66.7	720	2 J01676	ABI3 protein - Ara
29	38	66.7	1198	2 T43484	hypothetical prote

30	38	66.7	2411	2 A46299	tyrosine kinase su
31	37.5	65.8	203	2 AG3488	protein ybis precu
32	37.5	65.8	1520	2 T00273	hypothetical prote
33	37.5	65.8	5035	1 I46646	ryanodine receptor
34	37.5	65.8	5037	2 B35041	ryanodine receptor
35	37	64.9	13	2 A05174	tryptophyllin-13 -
36	37	64.9	134	2 D84672	hypothetical prote
37	37	64.9	142	2 C72523	hypothetical prote
38	37	64.9	211	2 A46458	human CRI homolog
39	37	64.9	218	2 A84688	hypothetical prote
40	37	64.9	251	2 T26856	hypothetical prote
41	37	64.9	264	2 E72642	hypothetical prote
42	37	64.9	273	2 AF2922	conserved hypoteth
43	37	64.9	287	2 G97696	hypothetical prote
44	37	64.9	288	2 T06282	alpha-gliadin prec
45	37	64.9	294	1 J02243	protein kinase [EC

ALIGNMENTS

RESULT 1

B45344
Probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)
C:Species: suid herpesvirus 1
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: B45344
R:Vicek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.
Virology 179, 365-377, 1990
A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented
A:Reference number: A45344; MUID:91021039; PMID:2171211
A:Accession: B45344
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1733 <VLC>
A:Cross-references: GB:M34651; NID:g334070; PIDN:AAA47471.1; PID:g334072
C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 77.2%; Score 44; DB 1; Length 1733;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 PRWPPP 7
| | | | |
Db 213 PRWPPP 218

RESULT 2

B40505
hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhuser or Becker)
C:Species: suid herpesvirus 1
C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 02-Sep-2000
C:Accession: B40505
R:Cheung, A.K.
J. Virol. 65, 5260-5271, 1991

A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies vir
A:Reference number: A40505; MUID:91374576; PMID:1654441
A:Accession: B40505
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1958 <CHE>
A:Cross-references: GB:M57505; NID:g334066; PIDN:AAA47468.1; PID:g334068
C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match 77.2%; Score 44; DB 2; Length 1958;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 PRWPPP 7
| | | | |
Db 423 PRWPPP 428

RESULT 3

C84681
 A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:BL1B22.50
 A:Experimental source: BAC clone BL1B22; strain OR74A
 C:Genetics:
 A:Gene: NCSP:BL1B22.50
 A:Map position: 6
 C:Superfamily: Neurospora crassa hypothetical protein BL1B22.50
 Query Match 71.9%; Score 41; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPP 6
 |||||
 Db 147 LPRWPP 152

RESULT 6
 B35670
 A:Cross-references: EC 2.7.1.112 2 - slime mold (Dictyostelium discoideum) (fra
 C:Species: Dictyostelium discoideum
 C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
 C:Accession: B35670
 R:Tan, J.L.; Spudich, J.A.
 Mol. Cell. Biol. 10, 3578-3583, 1990
 A:Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium dis
 A:Reference number: A35670; MUID:90287147; PMID:1972546
 A:Accession: B35670
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-410 <TAN>
 A:Cross-references: GB:M33784; NID:g167777; PIDN:AAA3203.1; PID:g167778
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:106-371/Domain: protein kinase homology <KIN>
 F:114-122/Region: protein kinase ATP-binding motif

Query Match 71.9%; Score 41; DB 2; Length 410;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPQL 9
 |||||
 Db 273 PRWRPPEL 280

RESULT 7
 A45617
 A:Cross-references: EC 2.7.1.11 - nematode (Haemonchus contortus)
 N:Alternate names: phosphofructokinase 1; phosphohexokinase
 C:Species: Haemonchus contortus
 C:Date: 03-Feb-1994 #sequence_revision 14-Jul-1994 #text_change 11-Jun-1999
 C:Accession: A45617
 R:Klein, R.D.; Olson, E.R.; Favreau, M.A.; Winterrowd, C.A.; Hatzenbuehler, N.T.; Shea
 Mol. Biochem. Parasitol. 48, 17-26, 1991
 A:Title: Cloning of a cDNA encoding phosphofructokinase from Haemonchus contortus.
 A:Reference number: A45617; MUID:92140452; PMID:1838137
 A:Accession: A45617
 A:Molecule type: mRNA
 A:Residues: 1-789 <KLE>
 A:Cross-references: GB:M59805; NID:g159184; PIDN:AAA29181.1; PID:g159185
 A:Experimental source: strain DF1020
 A:Note: This sequence is inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:79638, NCBI:P:79639)
 C:Superfamily: human 6-phosphofructokinase; 6-phosphofructokinase 1 homology
 C:Keywords: ATP; glycolysis; phosphoprotein; phosphotransferase
 F:34-342/Domain: 6-phosphofructokinase 1 homology <6PF1>
 F:421-708/Domain: 6-phosphofructokinase 1 homology <6PF2>

Query Match 71.9%; Score 41; DB 1; Length 789;
 Best Local Similarity 71.4%; Pred. No. 1.1e-02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 3

C84681
 A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:BL1B22.50
 A:Experimental source: BAC clone BL1B22; strain OR74A
 C:Genetics:
 A:Gene: NCSP:BL1B22.50
 A:Map position: 6
 C:Superfamily: Neurospora crassa hypothetical protein BL1B22.50
 Query Match 71.9%; Score 41; DB 2; Length 211;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPP 6
 |||||
 Db 147 LPRWPP 152

RESULT 6
 B35670
 A:Cross-references: EC 2.7.1.112 2 - slime mold (Dictyostelium discoideum) (fra
 C:Species: Dictyostelium discoideum
 C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
 C:Accession: B35670
 R:Tan, J.L.; Spudich, J.A.
 Mol. Cell. Biol. 10, 3578-3583, 1990
 A:Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium dis
 A:Reference number: A35670; MUID:90287147; PMID:1972546
 A:Accession: B35670
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-410 <TAN>
 A:Cross-references: GB:M33784; NID:g167777; PIDN:AAA3203.1; PID:g167778
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:106-371/Domain: protein kinase homology <KIN>
 F:114-122/Region: protein kinase ATP-binding motif

Query Match 71.9%; Score 41; DB 2; Length 410;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPQL 9
 |||||
 Db 273 PRWRPPEL 280

RESULT 4
 F75336
 A:Cross-references: EC 2.7.1.112 2 - slime mold (Dictyostelium discoideum) (fra
 C:Species: Dictyostelium discoideum
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: F75336
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75336
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-190 <WHI>
 A:Cross-references: GB:AE002032; GB:AE000513; NID:g6459715; PIDN:AAF11472.1; PID:g645970
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1919
 A:Map position: 1

Query Match 71.9%; Score 41; DB 2; Length 190;
 Best Local Similarity 83.3%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
 |||||
 Db 17 PKWPPP 22

RESULT 5
 T49803
 A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:BL1B22.50
 A:Experimental source: BAC clone BL1B22; strain OR74A
 C:Genetics:
 A:Gene: NCSP:BL1B22.50
 A:Map position: 6
 C:Superfamily: Neurospora crassa hypothetical protein BL1B22.50 [imported] - Neurospora crassa
 Query Match 71.9%; Score 41; DB 2; Length 190;
 Best Local Similarity 83.3%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
 |||||
 Db 17 PKWPPP 22

RESULT 5
 T49803
 A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:BL1B22.50
 A:Experimental source: BAC clone BL1B22; strain OR74A
 C:Genetics:
 A:Gene: NCSP:BL1B22.50
 A:Map position: 6
 C:Superfamily: Neurospora crassa hypothetical protein BL1B22.50 [imported] - Neurospora crassa
 Query Match 71.9%; Score 41; DB 2; Length 190;
 Best Local Similarity 83.3%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
 |||||
 Db 17 PKWPPP 22

RESULT 5
 T49803
 A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:BL1B22.50
 A:Experimental source: BAC clone BL1B22; strain OR74A
 C:Genetics:
 A:Gene: NCSP:BL1B22.50
 A:Map position: 6
 C:Superfamily: Neurospora crassa hypothetical protein BL1B22.50 [imported] - Neurospora crassa
 Query Match 71.9%; Score 41; DB 2; Length 190;
 Best Local Similarity 83.3%; Pred. No. 25;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
 |||||
 Db 17 PKWPPP 22

Qy 1 LPRWPPP 7
:|||||
Db 246 IPEWPPP 252

RESULT 8

isocitrate dehydrogenase homolog - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T03587
B:Umeda, M.; Hara, C.; Matsubayashi, Y.; Li, H.; Lui, Q.; Tadokoro, F.; Aotsuka, S.; Uchi
Plant Mol. Biol. 25, 469-478, 1994
A:Title: Expressed sequence tags from cultured cells of rice (*Oryza sativa* L.) under str
A:Reference number: 214962; MUID:94325471; PMID:8049371
A:Accession: T03587
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-61 <OMES>
A:Cross-references: EMBL:D21069; NID:9413915; PIDN:BAA04645.1; PID:9602777
A:Experimental source: callus
C:Genetics:
A:Note: AK88
C:Superfamily: yeast isocitrate dehydrogenase (NADPP+)

Query Match 70.2%; Score 40; DB 2; Length 61;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRWPPQL 9
:|||||
Db 15 PAWPPWL 22

RESULT 9

hypothetical protein ECS1526 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90819
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90819
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA834948.1; PID:q13360989; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECS1526

Query Match 70.2%; Score 40; DB 2; Length 166;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPPP 7
:|||||
Db 135 POWPPP 140

RESULT 10

hypothetical protein B0432.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T25460
R:Henkhaus, J.; Wohldmann, P.
Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid B0432.
A:Reference number: 220038

A:Accession: T25460
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-180 <HEN>
A:Cross-references: EMBL:U80836; PIDN:AAB37899.1; GSPDB:GN000020; CESP:B0432.3
A:Experimental source: strain Bristol N2; clone B0432
C:Genetics:
A:Gene: CESP:B0432.3
A:Map position: 2
A:Introns: 70/2; 106/2; 148/2
C:Superfamily: Caenorhabditis elegans hypothetical protein B0432.3

Query Match 70.2%; Score 40; DB 2; Length 180;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RWPPQL 9
:|||||
Db 148 RWPPPKM 154

RESULT 11

A87253
hypothetical protein CC0033 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87253
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87253
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <STO>
A:Cross-references: GB:AB005673; NID:g13421125; PIDN:AAK22021.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0033

Query Match 70.2%; Score 40; DB 2; Length 206;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPRWPPP 7
:|||||
Db 126 LPRWPEP 132

RESULT 12

G72754
probable phosphoenolpyruvate synthase APE0028 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: G72754
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: G72754
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <RAW>
A:Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA78937.1; PID:dl042713; PID:
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0028

Query Match 70.2%; Score 40; DB 2; Length 218;
Best Local Similarity 62.5%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPQ 8
:|||||
Db 64 IQRWPPPE 71

RESULT 13

T24543

hypothetical protein T05F1.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000

C:Accession: T24543

R:Burton, J.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19905

A:Accession: T24543

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-409 <WIL>

A:Cross-references: EMBL:Z81586; PIDN:CAB04697.1; GSPDB:GN00019; CESP:T05F1.8

A:Experimental source: clone T05F1

C:Genetics:

A:Gene: CESP:T05F1.8

A:Map position: 1

A:Introns: 337/3

C:Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology

Query Match 70.2%; Score 40; DB 2; Length 409;

Best Local Similarity 77.8%; Pred. No. 77;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPQ 9
||| |||||

Db 342 LPRPPQ 350

RESULT 14

T22888

hypothetical protein F58A3.1b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22888

R:Lennard, N.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19632

A:Accession: T22888

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-578 <WIL>

A:Cross-references: EMBL:Z81017; PIDN:CAB02672.1; GSPDB:GN00028; CESP:F58A3.1b

A:Experimental source: clone F58A3

C:Genetics:

A:Gene: CESP:F58A3.1b

A:Map position: X

A:Introns: 30/2; 66/3; 108/1; 209/3; 256/3; 285/1; 356/3

Query Match 70.2%; Score 40; DB 2; Length 578;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
|:||||

Db 566 POWPPP 571

RESULT 15

T22885

hypothetical protein F58A3.1a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22885

R:Lennard, N.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19632

A:Accession: T22885

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-592 <WIL>

A:Cross-references: EMBL:Z81017; PIDN:CAB02669.1; GSPDB:GN00028; CESP:F58A3.1a

A:Experimental source: clone F58A3

C:Genetics:

A:Gene: CESP:F58A3.1a

A:Map position: X

A:Introns: 44/2; 80/3; 122/1; 223/3; 270/3; 299/1; 370/3

Query Match 70.2%; Score 40; DB 2; Length 592;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
|:||||

Db 580 POWPPP 585

Search completed: November 26, 2002, 19:42:36

Job time: 10.9355 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 17:06:07 ; Search time 9.48387 Seconds
(without alignments)
126.452 Million cell updates/sec

Title: US-09-674-593-3
Perfect score: 57
Sequence: 1 LPRWPPQL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	9	21	AAV53810
2	57	100.0	84	21	AAV53809
3	45	78.9	104	22	AAO03578
4	45	78.9	195	22	AAU30711
5	44	77.2	90	22	ABG08009
6	44	77.2	443	22	AAU29704
7	44	77.2	709	19	AAW58988
8	44	77.2	709	22	AAW58968
9	44	77.2	1507	23	ABW97356
10	44	77.2	1958	15	AAW60620

11	43	75.4	53	23	ABP09141	Human ORFX protein
12	43	75.4	75	22	AAO07833	Human polypeptide
13	42.5	74.6	32	22	ABG19493	Novel human diagno
14	42	73.7	142	22	ABG02119	Novel human diagno
15	42	73.7	429	22	ABG19033	Novel human diagno
16	42	73.7	2424	22	ABW58924	Drosophila melanog
17	41	71.9	73	23	AAO21709	Human secreted pro
18	41	71.9	204	22	ABG27363	Novel human diagno
19	41	71.9	223	21	AAW41572	Human ORFX ORF1336
20	41	71.9	224	21	AAW51810	Gene 28 human secr
21	41	71.9	228	22	AAU51500	Propionibacterium
22	41	71.9	789	12	AAW15143	Haemochus contort
23	41	71.9	978	22	AAW65604	Novel protein kina
24	40	70.2	27	22	ABG15601	Novel human diagno
25	40	70.2	27	22	ABG15610	Novel human diagno
26	40	70.2	34	22	AAO13538	Human polypeptide
27	40	70.2	46	21	ABW32453	Human secreted pro
28	40	70.2	65	23	ABP08789	Human ORFX protein
29	40	70.2	79	22	ABG10176	Novel human diagno
30	40	70.2	108	22	AAU41561	Propionibacterium
31	40	70.2	285	22	AAW23514	Human EST encoded
32	40	70.2	336	21	AAW27989	Human secreted pro
33	40	70.2	336	21	AAW27990	Human secreted pro
34	40	70.2	336	22	AAW64535	Gene 41 human secr
35	40	70.2	336	22	AAW64536	Human secreted pro
36	40	70.2	398	22	ABG10177	Novel human diagno
37	40	70.2	457	22	ABW63679	Drosophila melanog
38	40	70.2	514	22	ABG10181	Novel human diagno
39	40	70.2	571	22	ABW67288	Drosophila melanog
40	40	70.2	585	20	AAW13375	Amino acid sequenc
41	40	70.2	654	22	AAU29036	Human PRO polypept
42	40	70.2	654	22	AAW06593	Human protein havi
43	40	70.2	654	22	AAW80243	Human PRO262 prote
44	40	70.2	796	22	AAW94811	Human protein sequ
45	39	68.4	43	22	AAW61771	Human brain expres

ALIGNMENTS

RESULT 1
AAV53810
ID AAV53810 standard; Peptide: 9 AA.
XX AC AAV53810;
XX DT 22-FEB-2000 (first entry)
XX DE Antigen encoded by tmour rejection antigen RUR-1 antisense cDNA.
XX DE Antisense; human; tumour rejection antigen; RUR-1; tumour; cancer;
XX KW renal cell carcinoma; colorectal carcinoma; melanoma; sarcoma;
XX KW leukaemia.
XX OS Homo sapiens.
XX OS WO9958546-A1.
XX PD 18-NOV-1999.
XX PF 13-MAY-1999; 99WO-US10424.
XX PR 13-MAY-1998; 98US-0085318.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Van Den Eynde B, Boon-Falleur T;
XX WPI: 2000-053076/04.
XX PT New isolated tumour rejection antigen RUR-1 nucleic acids, used for,
XX PT e.g. treatment of cancers -

PS Claim 17; Page 42; 75pp; English.

XX The present sequence represents an antigen derived from a protein
CC encoded by the antisense cDNA sequence of human tumour rejection
CC antigen RUR-1. The present antigen is recognised by CTL 361A/21.
CC The cDNA sequence is the antisense strand of a ubiquitously expressed
CC gene. The antisense strand codes for a polypeptide which is
CC preferentially expressed in tumour samples and tumour-derived cells
CC lines. The polypeptide is unrelated to any TRAP protein. The RUR-1
CC antisense cDNA sequence was isolated from a renal cell carcinoma line
CC LB9211-RCC. The RUR-1 nucleic acids and polypeptides can be used for
CC diagnosis, prognosis or treatment of a disorder characterized by the
CC expression of a RUR-1 antisense cDNA molecule or an expression product,
CC such as cancers, e.g. renal cell carcinoma, colorectal carcinoma,
CC melanoma, sarcoma or leukaemia.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 57; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPPQL 9
| | | | | | | | |
DB 1 LPRWPPQL 9

RESULT 2

AAV53809
ID AAV53809 standard; Protein; 84 AA.

AC AAV53809;

XX 22-FEB-2000 (first entry)

XX Protein encoded by tumour rejection antigen RUR-1 antisense cDNA.

XX Antisense; human; tumour rejection antigen; RUR-1; tumour; cancer;
KW renal cell carcinoma; colorectal carcinoma; melanoma; sarcoma;
KW leukaemia.

XX Homo sapiens.

XX WO9958546-A1.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US10424.

XX 13-MAY-1998; 98US-0085318.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Van Den Eynde B, Boon-Falleur T;

XX WPI; 2000-053076/04.

XX N-PSDB; AAZ36643.

XX New isolated tumour rejection antigen RUR-1 nucleic acids, used for,
PT e.g. treatment of cancers -

XX Claim 16; Fig 5; 75pp; English.

XX The present sequence is encoded by the antisense cDNA sequence of human
CC tumour rejection antigen RUR-1. This cDNA sequence is the antisense
CC strand of a ubiquitously expressed gene. The antisense strand codes for
CC a polypeptide which is preferentially expressed in tumour samples and
CC tumour-derived cells lines. The polypeptide is unrelated to any TRAP
CC protein. The cDNA sequence was isolated from a renal cell carcinoma line
CC LB9211-RCC. The RUR-1 nucleic acids and polypeptides can be used for
CC diagnosis, prognosis or treatment of a disorder characterized by the
CC expression of a RUR-1 antisense cDNA molecule or an expression product,
CC such as cancers, e.g. renal cell carcinoma, colorectal carcinoma,

CC melanoma, sarcoma or leukaemia.

XX SQ Sequence 84 AA;

Query Match 100.0%; Score 57; DB 21; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.43; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 LPRWPPQL 9

| | | | | | | | |
DB 38 LPRWPPQL 46

RESULT 3

AAO03578
ID AAO03578 standard; Protein; 104 AA.

XX AAO03578;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 17470.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI83509.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 17470; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 104 AA;

Query Match 78.9%; Score 45; DB 22; Length 104;

Best Local Similarity 85.7%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPPQ 8
I:|||||
Db 7 PQWPPQ 13

RESULT 4

AAU30711
ID AAU30711 standard; Protein; 195 AA.

XX AC AAU30711;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #1202.

XX DE Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US08656.

XX PR 18-APR-2000; 2000US-0552929.

XX PR 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy

PS Claim 20; Page 334; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX SQ Sequence 195 AA;

Query Match 78.9%; Score 45; DB 22; Length 195;

Best Local Similarity 85.7%; Pred. No. 44;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPPQ 8

I:|||||

Db 49 PRWPPR 55

RESULT 5

ABG08009

ID ABG08009 standard; Protein; 90 AA.

XX AC

ABG08009;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8000.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS72196.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

XX PS Claim 20; SEQ ID No 38368; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 90 AA;

Query Match 77.2%; Score 44; DB 22; Length 90;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPPP 7

I:|||||

Db 22 PRWPPP 27

RESULT 6

AAU29704

ID AAU29704 standard; Protein; 443 AA.

XX AC

AAU29704;

```

XX 18-DEC-2001 (first entry)
XX
XX
XX Novel human secreted protein #195.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US08656.
XX
XX 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
XX
XX Claim 20; Page 182-183; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. ARU29510-AAU3304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.
XX
XX Sequence 443 AA;
XX
XX Query Match 77.2%; Score 44; DB 22; Length 443;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 PRWPPP 7
XX | | | | |
XX Db 241 PRWPPP 246
XX
XX RESULT 7
XX AAW58988
XX ID AAW58988 standard; Protein; 709 AA.
XX
XX AC AAW58988;
XX
XX 11-SEP-1998 (first entry)
XX
XX Homo sapiens fetal kidney clone BD335_14 encoded protein.
XX
XX fetal; kidney; cDNA library; clone BD335_14; secreted protein
XX autoimmune disease; anti-inflammatory; immune; stimulation
XX suppression.

```

```

XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 255
XX /note= "undefined amino acid, encoded by AKC"
XX
XX WO9814576-A2.
XX
XX 09-APR-1998.
XX
XX 03-OCT-1997; 97WO-US18007.
XX
XX 04-OCT-1996; 96US-0726237.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Agostino MJ, Jacobs K, Lavallie ER, Mccoy JM, Marberg D;
XX Racie LA, Spaulding V, Treacy M;
XX
XX WPI; 1998-240082/21.
XX
XX N-PSDB; AAV11636.
XX
XX Nucleic acids encoding novel secreted proteins - useful as, e.g.
XX anti-inflammatory, immuno-stimulatory or suppressing agents
XX
XX Disclosure; Page 88-91; 110pp; English.
XX
XX The sequence is that of a secreted protein encoded by
XX an isolated polynucleotide which may be of use in the
XX production of therapeutic compositions for treating or
XX ameliorating a medical condition in a mammal. Such compositions
XX may be used for, e.g. research purposes as markers for
XX tissues, molecular weight markers for gels, primers or probes, for
XX nutrition as carbon, nitrogen or carbohydrate source. They can also be
XX used as a cytokine for cell proliferation and differentiation activity,
XX as immune stimulants or suppressors, e.g. for viral, bacterial or fungal
XX infections, for autoimmune diseases such as multiple sclerosis or
XX systemic lupus erythematosus, to regulate haematopoiesis, for tissue
XX growth, as an activator or inhibitor, or as a chemotactic or
XX chemokinetic, haemostatic and thrombocytic receptor/ligand,
XX anti-inflammatory or tumour inhibitor agents.
XX
XX Sequence 709 AA;
XX
XX Query Match 77.2%; Score 44; DB 19; Length 709;
XX Best Local Similarity 85.7%; Pred. No. 2e+02;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LPRWPPP 7
XX | | | | |
XX Db 135 LPQWPPP 141
XX
XX RESULT 8
XX AAB90668
XX ID AAB90668 standard; Protein; 709 AA.
XX
XX AC AAB90668;
XX
XX 07-JUN-2001 (first entry)
XX
XX Human BD335_14 protein sequence SEQ ID 5.
XX
XX Human; secreted protein; nutrient; cytokine modulator; proliferation;
XX differentiation; immune system modulator; tissue growth; chemotactic;
XX haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
XX haematopoiesis.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200119988-A1.
XX
XX

```

PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-US25135.
XX
PR 17-SEP-1999; 99US-0398829.
XX
PA (GEM) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;
XX
DR WPI; 2001-244801/25.
DR N-PSDB; AAF98376.
XX
PT Isolated nucleic acids encoding polypeptides, useful for modulating
PT e.g. cytokine and cell proliferation/differentiation activity, the
PT immune system and hematopoiesis regulating activity -
XX
PS Disclosure; Page 382-384; 557pp; English.
XX
CC Human cDNA clones represented in AAF98374 - AAF98489 encode secreted
CC proteins AAB90667 - AAB90750. The cDNA clones are isolated from various
CC tissue types, and may be used in the prevention, treatment and diagnosis
CC of diseases associated with inappropriate protein expression. The
CC polypeptides and nucleic acids may be used as nutrients or to modulate
CC cytokine and cell proliferation/differentiation activity and may also be
CC involved in modulation of the immune system. The cDNA sequences,
CC proteins, their agonists and/or antagonists exhibit hematopoiesis
CC regulating activity; tissue growth activity; activin/inhibin activity;
CC chemotactic/chemokinetic activity; haemostatic and thrombolytic
CC activity; receptor/ligand activity; anti-inflammatory activity;
CC hematopoiesis activity; cadherin/tumour suppressor activity; and/or
CC tumour inhibition activity. Included in the invention are probes
CC represented in AAF98490 - AAF98572 which are specific for the cDNA clones
CC encoding the secreted proteins.
XX
SQ Sequence 709 AA;
Query Match 77.2%; Score 44; DB 22; Length 709;
Best Local Similarity 85.7%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy . 1 LPRWPPP 7
Db 135 LPQWPPP 141
||:||||
RESULT 9
AAB97356
ID ABB97356 standard; Protein; 1507 AA.
XX
AC ABB97356;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 624.
XX
KW Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Dmanac RT;
XX
DR WPI; 2002-292408/33.
DR N-PSDB; ABN32542.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Example 2; SEQ ID NO 624; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 1507 AA;
Query Match 77.2%; Score 44; DB 23; Length 1507;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy . 1 LPRWPPP 7
Db 933 LPQWPPP 939
||:||||
RESULT 10
AAR60620
ID AAR60620 standard; Protein; 1958 AA.
XX
AC AAR60620;
XX
DT 15-MAY-1995 (first entry)
XX
DE Protein from ORF2 of Pseudorabies virus large latency transcript.
XX
KW Pseudorabies virus; PRV; LLV; large latency transcript;
KW attenuated virus; vaccine; early protein 0; EP0; HSV-1 ICP0;
KW protecting animals; deletion mutants; swine; ds.
XX
OS Pseudorabies virus.
XX
FH Key Location/Qualifiers
FH Region 330..333
FT /note="histidine rich region"
FT Region 363..368
FT /note="histidine rich region"
FT Region 381..384
FT /note="histidine rich region"
FT Region 547..573
FT /note="acidic residue rich region"
FT Region 1830..1833
FT /note="cysteine rich region"
FT Region 1837..1840
FT /note="cysteine rich region"
XX
PN US5352596-A.
XX
PD 04-OCT-1994.
XX
PF 11-SEP-1992; 92US-0945283.
XX
PR 11-SEP-1992; 92US-0945283.
XX
PA (USDA) US SEC OF AGRIC.

```

XX Cheung AK, Wesley RD;
PI
XX
XX
DR WPI: 1994-316187/39.
DR N-PSDB; AAO73500.
XX
XX New pseudorabies virus mutants for use in vaccine - having a
PT deletion and/or insertion in the early protein O gene or large
PT latency transcript gene
XX
XX Disclosure: Column 15-30; 43pp; English.
PS
XX
XX This protein sequence is encoded by ORF2 of LIT (large latency
CC transcript) from pseudorabies virus (PRV) as shown in AAO73500.
CC The basic sequence is derived from PRV strain InFh and PRV
CC strain Ka. The LIT overlaps and is transcribed in the opposite
CC orientation with respect to the EP0 (early polypeptide 0) and the
CC immediately early gene (IE180). EP0 is nonessential for replicatio,
CC LIT is the only gene expressed during PRV latency, and the IE180
CC gene is absolutely necessary for PRV replication. However there are
CC 2 copies of IE180 in the genome. It is expected that PRV lacking one
CC of the IE180 copies is viable. Deletions in the non-overlapping
CC regions of these 3 genes will generate single deletion routants,
CC while deletions in overlapping regions will generate double deletion
CC mutants. The invention is concerned with the construction of attenuated
CC viruses which have a reduced ability to reactivate from latency. This
CC can be achieved by functionally disabling the expression of the EP0
CC gene, or by disrupting the synthesis of the LIT, or both. (See also
CC AAO73501 and AARG0620-24)
XX
XX Sequence 1958 AA;
SQ
Query Match 77.2%; Score 44; DB 15; Length 1958;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 PRWPPP 7
DB 423 PRWPPP 428
RESULT 11
ID ABP09141 standard; Protein; 53 AA.
XX
XX AC ABP09141;
XX
XX 24-JUN-2002 (first entry)
XX
XX Human ORFX protein sequence SEQ ID NO:18264.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
XX Homo sapiens.
OS
XX WO200192523-A2.
PN
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US10836.
XX
XX 30-MAY-2000; 2000US-206132P.
XX
XX 29-AUG-2000; 2000US-228716P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX
XX Shimkets RA, Leach MD;
PI
XX
XX WPI: 2002-106308/14.
DR N-PSDB; ABN24893.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders
XX
XX Disclosure: SEQ ID 18264; 1037pp; English.
PS
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 53 AA;
SQ
Query Match 75.4%; Score 43; DB 23; Length 53;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 PRWPPP 8
DB 46 PSWPPPQ 52
RESULT 12
ID AAO07833 standard; Protein; 75 AA.
XX
XX AC AAO07833;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 21725.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX

```


XX -PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX PS
 DR WPI: 2001-514838/56.
 DR N-PSDB: AAI87764.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 FS Claim 20; SEQ ID NO 21725; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 75 AA;
 Query Match 75.4%; Score 43; DB 22; Length 75;
 Best Local Similarity 77.8%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LPRWPPQL 9
 Db 63 LPRWPPKAL 71
 RESULT 13
 ABG19493
 ID ABG19493 standard; Protein; 32 AA.
 XX AC ABG19493;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #19484.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 XX PD 30-MAR-2001; 2001WO-US08631.
 XX PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX PS WPI: 2001-639362/73.
 DR N-PSDB: AAS83680.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 20; SEQ ID NO 49852; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 32 AA;
 Query Match 74.6%; Score 42.5; DB 22; Length 32;
 Best Local Similarity 88.9%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 2 PR-WPPQL 9
 Db 13 PRWPPQL 21
 RESULT 14
 ABG02119
 ID ABG02119 standard; Protein; 142 AA.
 XX AC ABG02119;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #2110.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 XX PD 30-MAR-2001; 2001WO-US08631.
 XX PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX PS WPI: 2001-639362/73.
 DR N-PSDB: AAS66306.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID NO 32478; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX CC to restore normal activity of (II) or to treat disease states involving

XX CC (II). (II) is useful for generating antibodies against it, detecting or

XX CC quantitating a polypeptide in tissue, as molecular weight markers and as

XX CC a food supplement. (II) and its binding partners are useful in medical

XX CC imaging of sites expressing (II). (I) and (II) are useful for treating

XX CC disorders involving aberrant protein expression or biological activity.

XX CC The polypeptide and polynucleotide sequences have applications in

XX CC diagnostics, forensics, gene mapping, identification of mutations

XX CC responsible for genetic disorders or other traits to assess biodiversity

XX CC and to produce other types of data and products dependent on DNA and

XX CC amino acid sequences. ABG00010-ABG30377 represent novel human

XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 142 AA;

Query Match 73.7%; Score 42; DB 22; Length 142;

Best Local Similarity 85.7%; Pred. No. 85;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPRWPPP 7

Db 126 LPTWPPP 132

II IIII

RESULT 15

ABG19033

ID ABG19033 standard; Protein; 429 AA.

XX AC ABG19033;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #19024.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS83220.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity -

XX PS Claim 20; SEQ ID No 49392; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX CC to restore normal activity of (II) or to treat disease states involving

XX CC (II). (II) is useful for generating antibodies against it, detecting or

XX CC quantitating a polypeptide in tissue, as molecular weight markers and as

XX CC a food supplement. (II) and its binding partners are useful in medical

XX CC imaging of sites expressing (II). (I) and (II) are useful for treating

XX CC disorders involving aberrant protein expression or biological activity.

XX CC The polypeptide and polynucleotide sequences have applications in

XX CC diagnostics, forensics, gene mapping, identification of mutations

XX CC responsible for genetic disorders or other traits to assess biodiversity

XX CC and to produce other types of data and products dependent on DNA and

XX CC amino acid sequences. ABG00010-ABG30377 represent novel human

XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 429 AA;

Query Match 73.7%; Score 42; DB 22; Length 429;

Best Local Similarity 85.7%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPRWPPP 7

Db 96 LPAWPPP 102

II IIII

Search completed: November 26, 2002, 12:01:26

Job time : 14.4839 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 05:31:33 ; Search time 9.87097 Seconds
(without alignments)
187.856 Million cell updates/sec

Title: US-09-674-593-3
Perfect score: 57
Sequence: 1 LPRWPPPPQL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	84	Q9UBP8	Q9ubp8 homo sapien
2	44	77.2	156	Q9FTJ1	Q9ftj1 oryza sativ
3	44	77.2	767	Q96IR9	Q96ir9 homo sapien
4	44	77.2	1380	Q9P247	Q9p247 homo sapien
5	44	77.2	1958	Q9P340	Q9p340 pseudorabie
6	44	77.2	1995	Q8TF72	Q8tf72 homo sapien
7	43	75.4	642	Q9NSV4	Q9nsv4 caenorhabdi
8	42	73.7	211	Q9ZUV6	Q9zuw6 arabidopsis
9	42	73.7	769	Q52395	Q52395 mycobacteri
10	42	73.7	790	Q94GL4	Q94gl4 oryza sativ
11	42	73.7	2424	Q9VZ48	Q9vz48 drosophila
12	41	71.9	190	Q9RT47	Q9rt47 deincoccus
13	41	71.9	295	Q9SHP7	Q9shp7 arabidopsis
14	41	71.9	2785	Q8ZYB9	Q8zyb9 pyrobaculum
15	40	70.2	61	Q40658	Q40658 oryza sativ
16	40	70.2	166	Q8X310	Q8x310 escherichia

17	40	70.2	180	5	P90993	P90993 caenorhabdi
18	40	70.2	206	16	Q9AC33	Q9ac33 caulobacter
19	40	70.2	218	17	Q9YG73	Q9yg73 aeropyrum p
20	40	70.2	267	10	Q94DEL	Q94del oryza sativ
21	40	70.2	272	12	Q919W2	Q919m2 culex nigri
22	40	70.2	354	2	Q9REJ4	Q9rej4 bradyrhizob
23	40	70.2	409	5	Q9XUR1	Q9xur1 caenorhabdi
24	40	70.2	411	4	Q9BTQ4	Q9btq4 homo sapien
25	40	70.2	435	4	Q96DR0	Q96dr0 homo sapien
26	40	70.2	485	5	Q95RE0	Q95re0 drosophila
27	40	70.2	489	4	Q96E23	Q96e23 homo sapien
28	40	70.2	489	11	Q9BQC3	Q9bqc3 homo sapien
29	40	70.2	489	11	Q9CR25	Q9cr25 mus musculu
30	40	70.2	562	13	Q8UW75	Q8uw75 oryzias lat
31	40	70.2	571	5	O18357	O18357 drosophila
32	40	70.2	629	5	Q9U3E2	Q9u3e2 caenorhabdi
33	40	70.2	654	6	Q95LY1	Q95ly1 macaca fasc
34	40	70.2	654	6	Q95LV1	Q95lv1 macaca fasc
35	40	70.2	796	4	Q9H869	Q9h869 homo sapien
36	40	70.2	818	16	Q9F3E4	Q9f3e4 streptomyce
37	39	68.4	182	2	Q8RK85	Q8rkb5 fremyella d
38	39	68.4	221	10	Q9FR88	Q9fn88 arabidopsis
39	39	68.4	237	16	Q98ND6	Q98nd6 rhizobium l
40	39	68.4	274	4	Q96LW2	Q96lw2 homo sapien
41	39	68.4	299	16	Q8U8W0	Q8u8w0 agrobacteri
42	39	68.4	347	16	Q9K2U4	Q9kzu4 streptomyce
43	39	68.4	471	16	Q8X811	Q8xrl1 raistonia s
44	39	68.4	613	11	Q08841	Q08841 cavia porce
45	39	68.4	669	10	Q9ZWK0	Q9zwk0 arabidopsis

ALIGNMENTS

RESULT 1

Q9UBP8 PRELIMINARY: PRT; 84 AA.
ID Q9UBP8; AC Q9UBP8; DT 01-MAY-2000 (TREMBLrel. 13, Created)
RT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE RU2AS protein.
GN RU2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2006987; PubMed=10601354;
RA Van den Eynde B.J., Gaugler B., Probst-Keppler M., Michaux L.,
RA Devuyt O., Lorge F., Weynants P., Boon T.;
RT "A new antigen recognized by cytolytic T lymphocytes on a human kidney
RL tumor results from reverse strand transcription."
RL J. Exp. Med. 190:1793-1800(1999).
DR EMBL; AF181722; AAF23613.1;
DR EMBL; AF181720; AAF23611.1;
SQ SEQUENCE 84 AA; 8969 MW; 70B739F173A9E560 CRC64;

Query Match 100.0%; Score 57; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPPPPQL 9
Db 38 LPRWPPPPQL 46

RESULT 2

Q9FTJ1 PRELIMINARY: PRT; 156 AA.
ID Q9FTJ1; AC Q9FTJ1; DT 01-MAR-2001 (TREMBLrel. 16, Created)

```
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
GN P0410E01.9 protein.
OS Oryza sativa (rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
   clone: P0410E01."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002866; BAB17088.1;
SQ SEQUENCE 156 AA; 16741 MW; 6677477714F4E3B6 CRC64;

Query Match 77.2%; Score 44; DB 10; Length 156;
Best Local Similarity 77.8%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPRWPPPL 9
DB 74 LPPWPPPL 82

RESULT 3
Q96IR9 ID Q96IR9 PRELIMINARY; PRT; 767 AA.
AC Q96IR9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to shroom (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC007291; AAH07291.1;
FT NON_TER 1
SQ SEQUENCE 767 AA; 83291 MW; 1BF2520CD38C02B5 CRC64;

Query Match 77.2%; Score 44; DB 4; Length 767;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPPP 7
DB 193 LPQWPPP 199

RESULT 4
Q9P247 ID Q9P247 PRELIMINARY; PRT; 1380 AA.
AC Q9P247;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1481 protein (fragment).
GN KIAA1481.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 77.2%; Score 44; DB 4; Length 1995;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
DB 423 PRWPPP 428

RESULT 6
Q8TF72 ID Q8TF72 PRELIMINARY; PRT; 1995 AA.
AC Q8TF72;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Shroom-related protein.
GN HSHRML.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Koide M., Iio A., Obata K., Inagaki M., Yokota M., Ono T., Tuan R.S.;
RT "Molecular cloning of FKBP12-associated protein."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB055660; BAB84689.1;
SQ SEQUENCE 1995 AA; 216660 MW; D0914F4A034BAADA CRC64;
```

```
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohata O.;
RT "Prediction of the coding sequences of unidentified human
   genes. XVII. The complete sequences of 100 new cDNA clones from brain
   which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
DR EMBL: AB040914; BAA96005.1;
FT NON_TER 1
SQ SEQUENCE 1380 AA; 149753 MW; 973465CFFB9BC458 CRC64;

Query Match 77.2%; Score 44; DB 4; Length 1380;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPPP 7
DB 806 LPQWPPP 812

RESULT 5
Q69340 ID Q69340 PRELIMINARY; PRT; 1958 AA.
AC Q69340;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF1, ORF2, and ORF3.
OS Pseudorabies virus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIANA-FUNKHAUSER;
RX MEDLINE=91374576; PubMed=1654441;
RA Cheung A.K.;
RT "Cloning of the latency gene and the early protein 0 gene of
   pseudorabies virus."
RL J. Virol. 65:5260-5271(1991).
DR EMBL: M57505; AAA47468.1;
SQ SEQUENCE 1958 AA; 193474 MW; 0DEC36EDC29E2E68 CRC64;

Query Match 77.2%; Score 44; DB 12; Length 1958;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPP 7
DB 423 PRWPPP 428

RESULT 6
Q8TF72 ID Q8TF72 PRELIMINARY; PRT; 1995 AA.
AC Q8TF72;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Shroom-related protein.
GN HSHRML.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Koide M., Iio A., Obata K., Inagaki M., Yokota M., Ono T., Tuan R.S.;
RT "Molecular cloning of FKBP12-associated protein."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB055660; BAB84689.1;
SQ SEQUENCE 1995 AA; 216660 MW; D0914F4A034BAADA CRC64;
```

```

Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPRWPPPP 7
DB 1421 LPQWPPPP 1427

RESULT 7
Q9NSV4
ID Q9NSV4 PRELIMINARY; PRT; 642 AA.
AC Q9NSV4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 70.4 kDa protein.
GN C53D5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Murray J., Rohlfing T., Morris M.;
RT "The sequence of C. elegans cosmid C53D5.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006624; AAF39787.1; -.
DR InterPro; IPR000101; Gglutnsptase.
DR Pfam; PF01019; G-glu_transp; 1.
DR PRINTS; PR01210; GGTTRANSPTASE.
DR TIGRFAMS; TIGR00066; g_glu_trans; 1.
KW Hypothetical protein.
SQ SEQUENCE 642 AA; 70425 MW; 9D7601AB07110F3C CRC64;

Query Match 75.4%; Score 43; DB 5; Length 642;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPRWPPPPQ 9
DB 87 LPRWPGPSL 95

RESULT 8
Q9ZUV6
ID Q9ZUV6 PRELIMINARY; PRT; 211 AA.
AC Q9ZUV6;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Atg28140 protein.
GN AT2G28140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005851; AAC98452.1; -.
SQ SEQUENCE 211 AA; 24340 MW; B7DDF07BAC65BB89 CRC64;

Query Match 73.7%; Score 42; DB 10; Length 211;
Best Local Similarity 72.7%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 LPRW--PPPPQ 9
DB 156 LPQWRHPPPPQ 166

RESULT 9
O52395
ID O52395 PRELIMINARY; PRT; 769 AA.
AC O52395;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cell division protein.
GN FTSH.
OS Mycobacterium smegmatis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SN2;
RA Anilkumar G., Ajitkumar P.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF037269; AAC32257.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR000642; Peptidase_M41.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF01434; Peptidase_M41; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMS; TIGR01241; FlsH_fam; 1.
DR PROSITE; PS00674; AAA; 1.
KW ATP-binding.
SQ SEQUENCE 769 AA; 83649 MW; 738F37A17C78178E CRC64;

Query Match 73.7%; Score 42; DB 2; Length 769;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPPPQ 8
DB 695 PGWPPPPQ 701

RESULT 10
Q94GL4
ID Q94GL4 PRELIMINARY; PRT; 790 AA.

```

```

AC Q94GL4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA de la Bastide M., Nascimento L., Spiegel T., Kirchoff K., Preston R.,
RA King L., Vil M.D., Baker J., Zutavern T., Santos L., Bell M.,
RA Miller B., Kuit K., Rodriguez S., Cunniss D.M., Balija V., Shah R.,
RA Bahret A., Bal H., Palmer L., Yang C., O'Shaughnessy A., Dedhia N.,
RA McCombie W.R.;
RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB00044114, from Chromosome 10."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC090487; AAK92665.1;
SQ SEQUENCE 790 AA; 87798 MW; 245EB1FEEB28D55D CRC64;

Query Match 73.7%; Score 42; DB 10; Length 790;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPRWPPQL 9
Db '400 VPPWPPPSL 408

RESULT 11
Q9VZ48
ID Q9VZ48 PRELIMINARY; PRT; 2424 AA.
AC Q9VZ48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG2174 protein.
GN CG2174.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Fink M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003484; AAF47980.1;
DR HSSP; P08799; LMND.
DR FlyBase; FBgn0030252; CG2174.
DR InterPro; IPR000048; IQ-region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR000857; MYTH4.
DR Pfam; PF00612; IQ; 3.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00784; MYTH4; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR SMART; SM00139; MYTH4; 1.
SQ SEQUENCE 2424 AA; 267616 MW; 8AD62AA33F9AA5D0 CRC64;

Query Match 73.7%; Score 42; DB 5; Length 2424;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RWPAPPQ 8
Db 1714 RWPAPPQ 1719

RESULT 12
Q9RT47
ID Q9RT47 PRELIMINARY; PRT; 190 AA.
AC Q9RT47;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DR1919.
GN DR1919.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Yamashiro J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarewicz K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE002031; AAF11472.1;
DR TIGR; DR1919;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 190 AA; 21327 MW; 30326C4982CC4721 CRC64;

```

Query Match 71.9%; Score 41; DB 16; Length 190;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPP 7
 Db 17 PKWPP 22

RESULT 13
 Q9SH7 PRELIMINARY; PRT; 295 AA.
 AC Q9SH7;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE FK23.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Shinn P., Brooks S., Buehler E., Dunn P., Khan S., Kim C.,
 Walker M., Brooks S., Altafi H., Araujo R., Conn L., Conway A.B.,
 RA Gonzalez A., Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J.,
 RA Liu S., Lueros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V.,
 RA Yu G., Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC FK23 from chromosome
 I.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007508; AAG22834.1; -;
 DR InterPro; IP8001087; Lipase_GDSL.
 SQ SEQUENCE 295 AA; 32918 MW; 82F3ED7153BEB2EC CRC64;

Query Match 71.9%; Score 41; DB 10; Length 295;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPRWP 6
 Db .56 LPRWP 61

RESULT 14
 Q82YB9 PRELIMINARY; PRT; 2785 AA.
 AC Q82YB9;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE PARE2b.
 GN PAE0850.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792869;
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
 DR EMBL; AE009790; AAL63076.1; -;
 KW Complete proteome.
 SQ SEQUENCE 2785 AA; 300989 MW; 1E144809E1D59C2D CRC64;

Query Match 71.9%; Score 41; DB 17; Length 2785;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 PRWPPQ 8
 Db 74 PRWPPQ 80

RESULT 15
 Q40658 PRELIMINARY; PRT; 61 AA.
 AC Q40658;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)
 DE Isocitrate dehydrogenase (Fragment).
 GN AK88.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94323471; PubMed=8049371;
 RA Umeda M., Hara C., Matsubayashi Y., Li H., Lui Q., Tadokoro F.,
 RA Aotsuka S., Uchimiya H.;
 RT "Expressed sequence tags from cultured cells of rice (Oryza sativa L.)
 under stressed conditions: analysis of transcripts of genes engaged in
 ATP-generating pathways.";
 RL Plant Mol. Biol. 25:469-478 (1994).
 DR EMBL; D21069; BAA04645.1; -;
 FT NON_TER 1 1
 FT NON_TER 61 61
 SQ SEQUENCE 61 AA; 7204 MW; CA1655D36E56A1DE CRC64;

Query Match 70.2%; Score 40; DB 10; Length 61;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PRWPPQL 9
 Db 15 PAWPPWL 22

Search completed: November 26, 2002, 17:59:23
 Job time : 16.871 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 12:00:06 ; Search time 3.19355 Seconds
(without alignments)
82.919 Million cell updates/sec

Title: US-09-674-593-3
Perfect score: 57
Sequence: 1 LPRWPPQL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	77.2	1958	1	US-07-945-283-2
2	39	68.4	78	1	US-08-487-359-5
3	39	68.4	78	1	US-08-222-798A-5
4	39	68.4	1481	2	US-08-616-844-40
5	39	68.4	1481	2	US-08-599-654-40
6	39	68.4	1481	3	US-08-944-868A-40
7	39	68.4	1481	3	US-08-944-423A-40
8	39	68.4	1481	3	US-08-944-496-40
9	38	66.7	492	4	US-09-724-864-39
10	37	64.9	10	1	US-08-255-272-1
11	37	64.9	357	5	PCT-US91-00899-14
12	37	64.9	369	2	US-08-424-224-2
13	37	64.9	369	2	PCT-US94-02891-69
14	37	64.9	405	1	US-07-914-281-8
15	37	64.9	405	1	US-08-393-246-8
16	37	64.9	405	1	US-08-525-058A-8
17	37	64.9	405	2	US-08-483-151-4
18	37	64.9	405	2	US-08-696-731-8
19	37	64.9	405	4	US-09-042-531-8
20	37	64.9	577	4	US-09-413-814-4
21	37	64.9	1001	4	US-08-884-569A-2
22	37	64.9	1130	2	US-08-519-547A-6
23	36	63.2	20	4	US-08-602-999A-128
24	36	63.2	20	4	US-08-630-915A-181
25	36	63.2	20	4	US-08-630-915A-223
26	36	63.2	20	4	US-09-500-124-128
27	36	63.2	24	4	US-08-630-915A-41

28	36	63.2	79	1	US-08-487-359-2	Sequence 2, Appli
29	36	63.2	79	1	US-08-487-359-6	Sequence 6, Appli
30	36	63.2	79	1	US-08-222-798A-2	Sequence 2, Appli
31	36	63.2	79	1	US-08-222-798A-6	Sequence 6, Appli
32	36	63.2	170	4	US-08-259-451-19	Sequence 19, Appli
33	36	63.2	356	4	US-09-092-315-12	Sequence 12, Appli
34	36	63.2	393	4	US-09-390-131-8	Sequence 2, Appli
35	36	63.2	2289	3	US-09-051-019-2	Sequence 2, Appli
36	35.5	62.3	129	4	US-09-199-637A-97	Sequence 97, Appli
37	35	61.4	13	4	US-08-602-999A-80	Sequence 80, Appli
38	35	61.4	13	4	US-08-278-865-80	Sequence 80, Appli
39	35	61.4	13	4	US-09-500-124-80	Sequence 80, Appli
40	35	61.4	15	4	US-08-602-999A-342	Sequence 342, App
41	35	61.4	15	4	US-09-500-124-342	Sequence 342, App
42	35	61.4	45	4	US-08-602-999A-48	Sequence 48, Appli
43	35	61.4	45	4	US-08-278-865-48	Sequence 48, Appli
44	35	61.4	45	4	US-09-500-124-48	Sequence 48, Appli
45	35	61.4	80	4	US-09-134-001C-2843	Sequence 2843, Ap

ALIGNMENTS

RESULT 1
US-07-945-283-2
Sequence 2, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The EP0 and LLT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/945,283
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1958 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-945-283-2

Query Match 77.2%; Score 44; DB 1; Length 1958;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPPP 7
Db 423 PRWPPP 428

RESULT 2

US-08-487-359-5
; Sequence 5, Application US/08487359
; Patent No. 5633229
; GENERAL INFORMATION:
; APPLICANT: KOKRYAKOV, VLADIMIR N.
; APPLICANT: HARMIG, SYLVIA S.L.
; APPLICANT: LEHRER, ROBERT I.
; TITLE OF INVENTION: PROPHENINS-ANTIBIOTIC PEPTIDES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; STATE: CA
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,359
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/222,798
; FILING DATE: 05-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2000-0553.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-487-359-5

Query Match 68.4%; Score 39; DB 1; Length 78;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPQ 8
Db 30 PRFPQQ 36

RESULT 3
US-08-222-798A-5
; Sequence 5, Application US/08222798A
; Patent No. 5804553
; GENERAL INFORMATION:
; APPLICANT: KOKRYAKOV, VLADIMIR N.
; APPLICANT: HARMIG, SYLVIA S.L.
; APPLICANT: LEHRER, ROBERT I.
; TITLE OF INVENTION: PROPHENINS-ANTIBIOTIC PEPTIDES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; STATE: CA
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,798A
; FILING DATE: 05-APR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2000-0553.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-222-798A-5

Query Match 68.4%; Score 39; DB 1; Length 78;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWPPQ 8
Db 30 PRFPQQ 36

RESULT 4
US-08-616-844-40
; Sequence 40, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,844
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-616-844-40

Query Match 68.4%; Score 39; DB 2; Length 1481;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPPOL 9
| | | | |
Db 8 RWPPPLL 14

RESULT 5

US-08-599-654-40
; Sequence 40, Application US/08599654
; Patent No. 5882925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,654

; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573

; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844

; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A

; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-041
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864

; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown

; MOLECULE TYPE: protein
US-08-599-654-40

Query Match 68.4%; Score 39; DB 2; Length 1481;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPPOL 9
| | | | |
Db 8 RWPPPLL 14

RESULT 6

US-08-944-868A-40
; Sequence 40, Application US/08944868A
; Patent No. 6018025
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,868A

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,654

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844

; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A

; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864

; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown

; MOLECULE TYPE: protein
US-08-944-868A-40

Query Match 68.4%; Score 39; DB 3; Length 1481;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPPOL 9
| | | | |
Db 8 RWPPPLL 14

RESULT 7

US-08-944-423A-40
; Sequence 40, Application US/08944423A
; Patent No. 6020463
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,423A
; FILING DATE: 06-OCT-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: JUN-07-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-944-423A-40

Query Match 68.4%; Score 39; DB 3; Length 1481;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWPPPOL 9
Db 8 RWPPPLL 14

RESULT 8
US-08-944-496-40
; Sequence 40, Application US/08944496
; Patent No. 612433
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,496
; FILING DATE: 06-OCT-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654

; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-944-496-40

Query Match 68.4%; Score 39; DB 3; Length 1481;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWPPPOL 9
Db 8 RWPPPLL 14

RESULT 9
US-09-724-864-39
; Sequence 39, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.1050U1
; CURRENT APPLICATION NUMBER: US/09/724,864
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Mouse
; US-09-724-864-39

Query Match 66.7%; Score 38; DB 4; Length 492;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
Db 425 PTWPPP 430

RESULT 10
US-08-253-272-1
; Sequence 1, Application US/08255272
; Patent No. 5627268
; GENERAL INFORMATION:
; APPLICANT: Kumar, Ramesh
; APPLICANT: Sharma, Ajay
; APPLICANT: Khoury-Christianson, Anastasia
; APPLICANT: M.

;; TITLE OF INVENTION: Production of Therapeutic Peptides in
;; TITLE OF INVENTION: Transgenic Animals as a Fusion with Hemoglobin
;; NUMBER OF SEQUENCES: 32
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PENNIE & EDMONDS
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/255,272
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A.
;; REGISTRATION NUMBER: 30742
;; REFERENCE/DOCKET NUMBER: 6794-032
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; US-08-255-272-1

Query Match 64.9%; Score 37; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 6.4;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRWPPQL 9
Db 1 PWRPQI 8

RESULT 11
PCT-US91-00899-14
;; Sequence 14, Application PC/TUS9100899
;; GENERAL INFORMATION:
;; APPLICANT: Lowe, John B.
;; TITLE OF INVENTION: Method and Products For the Synthesis of
;; TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,
;; TITLE OF INVENTION: or as Free Molecules, and for the Isolation of Cloned
;; TITLE OF INVENTION: Genetic Sequences That Determine These Structures
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,
;; ADDRESSEE: P.C.
;; STREET: 1755 Jefferson Davis Highway, Suite 400
;; CITY: Arlington
;; STATE: Virginia
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US91/00899
;; FILING DATE: 19910214
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lavalleye Ph.D., Jean-Paul

;; REGISTRATION NUMBER: 31,451
;; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)521-5940
;; TELEFAX: (703)486-2347
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 357 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: C-terminal
;; PCT-US91-00899-14

Query Match 64.9%; Score 37; DB 5; Length 357;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
Db 71 PDWPPP 76

RESULT 12
US-08-424-224-2
;; Sequence 2, Application US/08424224
;; Patent No. 5912173
;; GENERAL INFORMATION:
;; APPLICANT: LEONARD, WARREN J.
;; TITLE OF INVENTION: MURINE IL-2R CDNA AND
;; TITLE OF INVENTION: USES THEREOF
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN
;; STREET: 345 PARK AVE.
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORD PERFECT # 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,224
;; FILING DATE:
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/121,435
;; FILING DATE: 14-SEPT-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WILLIAM S. FEILER
;; REGISTRATION NUMBER: 26,728
;; REFERENCE/DOCKET NUMBER: 2026-4061US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; TELEX: 421792
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 369
;; TYPE: AMINO ACID
;; TOPOLOGY: UNKNOWN
;; MOLECULE TYPE:
;; DESCRIPTION: PROTEIN
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: MURINE
;; INDIVIDUAL ISOLATE: IL-2R
;; US-08-424-224-2

Query Match 64.9%; Score 37; DB 2; Length 369;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
DB 356 PYWPPP 361

RESULT 13

PCT-US94-02891-69

; Sequence 69, Application PC/TUS9402891
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN
; APPLICANT: SERVICES
; APPLICANT: OFFICE OF TECHNOLOGY TRANSFER, NATIONAL
; APPLICANT: INSTITUTES OF HEALTH, BOX 017, BETHESDA, MARYLAND 20892 USA
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: XSCID
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT # 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02891
; FILING DATE:

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/031,143
; FILING DATE: 12-MAR-1993
; APPLICATION NUMBER: 08/121,435
; FILING DATE: 14-SEPT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792

; INFORMATION FOR SEQ. ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 369

TYPE: AMINO ACID

TOPOLOGY: UNKNOWN

MOLECULE TYPE:

DESCRIPTION: PROTEIN

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: MURINE

INDIVIDUAL ISOLATE: IL-2R

PCT-US94-02891-69

Query Match 64.9%; Score 37; DB 5; Length 369;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
DB 356 PYWPPP 361

RESULT 14

US-07-914-281-8
; Sequence 8, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 405 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-07-914-281-8

Query Match 64.9%; Score 37; DB 1; Length 405;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
DB 122 PDWPPP 127

RESULT 15

US-08-393-246-8

; Sequence 8, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.

; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

; COMPUTER READABLE FORM:

FLOPPY disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/393,246
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 405 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-393-246-8

Query Match 64.9%; Score 37; DB 1; Length 405;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
Db 122 PDNPPP 127

Search completed: November 26, 2002, 19:43:20
Job time : 5.19355 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 12:01:36 ; Search time 4.74194 Seconds
(without alignments)
30.223 Million cell updates/sec

Title: US-09-674-593-3

Perfect score: 57

Sequence: 1 LPRWPPQL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	70.2	654	9	US-09-905-291A-177
2	40	70.2	654	10	US-09-909-320-177
3	40	70.2	654	10	US-09-909-088A-177
4	40	70.2	654	12	US-10-052-586-26
5	39	68.4	43	10	US-09-864-761-44755
6	39	68.4	229	9	US-09-764-868-831
7	39	68.4	1481	10	US-09-371-900-40
8	39	68.4	1481	10	US-09-924-417-60
9	38	66.7	71	10	US-09-864-761-42815
10	38	66.7	348	10	US-09-764-898-188
11	38	66.7	482	10	US-09-764-853-568
12	38	66.7	1198	10	US-09-866-582-36
13	38	66.7	1241	12	US-10-001-215-5
14	37	64.9	39	10	US-09-864-761-46140
15	37	64.9	52	10	US-09-864-761-40064
16	37	64.9	79	10	US-09-925-300-1002
17	37	64.9	83	10	US-09-790-264-17
18	37	64.9	103	10	US-09-764-864-1448
19	37	64.9	160	10	US-09-764-864-1020

20	37	64.9	184	10	US-09-925-297-553
21	37	64.9	210	10	US-09-790-264-23
22	37	64.9	247	10	US-09-790-264-20
23	37	64.9	252	10	US-09-738-396-4
24	37	64.9	252	10	US-09-733-774A-4
25	37	64.9	252	10	US-09-771-961-4
26	37	64.9	287	10	US-09-764-898-218
27	37	64.9	288	10	US-09-764-853-841
28	37	64.9	288	10	US-09-764-898-290
29	37	64.9	314	10	US-09-764-853-670
30	37	64.9	327	10	US-09-738-396-2
31	37	64.9	327	10	US-09-773-774A-2
32	37	64.9	327	10	US-09-771-961-2
33	37	64.9	357	9	US-09-736-457-1817
34	37	64.9	357	9	US-09-902-941-1817
35	37	64.9	357	10	US-09-833-790-436
36	37	64.9	359	10	US-09-925-301-1127
37	37	64.9	369	9	US-09-895-593-12
38	37	64.9	369	10	US-09-895-943-12
39	37	64.9	393	9	US-09-764-868-690
40	37	64.9	405	10	US-09-863-475A-8
41	37	64.9	830	10	US-09-064-199-19
42	37	64.9	1106	10	US-09-064-199-17
43	37	64.9	1130	9	US-10-104-595-6
44	37	64.9	1130	10	US-09-064-199-18
45	37	64.9	1207	10	US-09-064-199-16

ALIGNMENTS

RESULT 1

US-09-905-291A-177
; Sequence 177, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222

```

; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 177
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-177

```

Query Match 70.2%; Score 40; DB 9; Length 654;

Best Local Similarity 75.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 LPRWPPQ 8
    |||||
Db 526 LPKWYPQ 533

```

RESULT 2

```

US-09-909-320-177
; Sequence 177, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

```

```

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 177
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-177

```

Query Match 70.2%; Score 40; DB 10; Length 654;

Best Local Similarity 75.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 LPRWPPQ 8
    |||||
Db 526 LPKWYPQ 533

```

RESULT 3

```

US-09-909-088B-177
; Sequence 177, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.

```

APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2001-07-18
PRIORITY APPLICATION NUMBER: US 60/143,048
PRIORITY FILING DATE: 1999-07-07
PRIORITY APPLICATION NUMBER: US 60/145,698
PRIORITY FILING DATE: 1999-07-26
PRIORITY APPLICATION NUMBER: US 60/146,222
PRIORITY FILING DATE: 1999-07-28
PRIORITY APPLICATION NUMBER: PCT/US99/20594
PRIORITY FILING DATE: 1999-09-08
PRIORITY APPLICATION NUMBER: PCT/US99/20944
PRIORITY FILING DATE: 1999-09-13
PRIORITY APPLICATION NUMBER: PCT/US99/21090
PRIORITY FILING DATE: 1999-09-15
PRIORITY APPLICATION NUMBER: PCT/US99/21547
PRIORITY FILING DATE: 1999-09-15
PRIORITY APPLICATION NUMBER: PCT/US99/23089
PRIORITY FILING DATE: 1999-10-05
PRIORITY APPLICATION NUMBER: PCT/US99/28214
PRIORITY FILING DATE: 1999-11-29
PRIORITY APPLICATION NUMBER: PCT/US99/28313
PRIORITY FILING DATE: 1999-11-30
PRIORITY APPLICATION NUMBER: PCT/US99/28564
PRIORITY FILING DATE: 1999-12-02
PRIORITY APPLICATION NUMBER: PCT/US99/28565
PRIORITY FILING DATE: 1999-12-02
PRIORITY APPLICATION NUMBER: PCT/US99/30095
PRIORITY FILING DATE: 1999-12-16
PRIORITY APPLICATION NUMBER: PCT/US99/30911
PRIORITY FILING DATE: 1999-12-20
PRIORITY APPLICATION NUMBER: PCT/US99/30999
PRIORITY FILING DATE: 1999-12-20
PRIORITY APPLICATION NUMBER: PCT/US00/00219
PRIORITY FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 177
LENGTH: 654
TYPE: PRT
ORGANISM: Homo sapiens
US-09-909-088B-177

Query Match 70.2%; Score 40; DB 10; Length 654;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPQ 8
||:|:|
Db 526 LPKWPFQ 533

RESULT 4
US-10-052-586-26
Sequence 26, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: F3430R1C1
CURRENT FILING DATE: 2002-01-15
PRIORITY APPLICATION NUMBER: US/10/052,586
PRIORITY FILING DATE: 1997-09-18
PRIORITY APPLICATION NUMBER: 60/059263
PRIORITY FILING DATE: 1997-09-18
PRIORITY APPLICATION NUMBER: 60/059266
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/062250
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063120
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063121
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063486
PRIORITY FILING DATE: 1997-10-21
PRIORITY APPLICATION NUMBER: 60/063540
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063541
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063544
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063564
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063734
PRIORITY FILING DATE: 1997-10-29
PRIORITY APPLICATION NUMBER: 60/063870
PRIORITY FILING DATE: 1997-10-31
PRIORITY APPLICATION NUMBER: 60/064103
PRIORITY FILING DATE: 1997-10-31
PRIORITY APPLICATION NUMBER: 60/065311
PRIORITY FILING DATE: 1997-11-13
PRIORITY APPLICATION NUMBER: 60/066120
PRIORITY FILING DATE: 1997-11-21
PRIORITY APPLICATION NUMBER: 60/066466
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/066772
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/069335
PRIORITY FILING DATE: 1997-12-11
PRIORITY APPLICATION NUMBER: 60/069425
PRIORITY FILING DATE: 1997-12-12
PRIORITY APPLICATION NUMBER: 60/069870
PRIORITY FILING DATE: 1997-12-17
PRIORITY APPLICATION NUMBER: 60/068017
PRIORITY FILING DATE: 1997-12-18
PRIORITY APPLICATION NUMBER: 60/077450
PRIORITY FILING DATE: 1998-03-10
PRIORITY APPLICATION NUMBER: 60/077632
PRIORITY FILING DATE: 1998-03-11
PRIORITY APPLICATION NUMBER: 60/077649
PRIORITY FILING DATE: 1998-03-11
PRIORITY APPLICATION NUMBER: 60/078886
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 60/078939
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 60/079664
PRIORITY FILING DATE: 1998-03-27
PRIORITY APPLICATION NUMBER: 60/079786
PRIORITY FILING DATE: 1998-03-27
PRIORITY APPLICATION NUMBER: 60/080107
PRIORITY FILING DATE: 1998-03-31

;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029

;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 70.2%; Score 40; DB 12; Length 654;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Labels 0; Gaps 0;

Qy 1 LPRWPPQ 8
Db 526 LPKWPYPQ 533

RESULT 5

US-09-864-761-44755
; Sequence 44755, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44755
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC019194.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.64
; OTHER INFORMATION: EST HUMAN HIT: AA283201.1, EVALUATE 7.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O27719, EVALUATE 1.20e+00
US-09-864-761-44755

Query Match 68.4%; Score 39; DB 10; Length 43;
Best Local Similarity 85.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPRWPP 7
| | | | |
Db 9 LPRWGP 15

RESULT 6
US-09-764-868-831
; Sequence 831, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 831
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (125)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-831

Query Match 68.4%; Score 39; DB 9; Length 229;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LPRWPPQ 8
| | | | |
Db 32 LPEWVPQ 39

RESULT 7

US-09-371-900-40
; Sequence 40, Application US/09371900
; Patent No. US20020137700A1
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/371,900
; FILING DATE: 11-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-104
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-371-900-40

Query Match 68.4%; Score 39; DB 10; Length 1481;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPPOL 9
Db 8 RWPPPLL 14

RESULT 8

US-09-924-417-60
Sequence 60, Application US/09924417
Patent No. US2002014241A1
GENERAL INFORMATION:
APPLICANT: Falb, Dean
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
DISEASE

NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/924,417
FILING DATE: 07-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,286
FILING DATE: 04-MAR-1998
APPLICATION NUMBER: 08/870,434
FILING DATE: 06-JUN-1997
APPLICATION NUMBER: 08/799,910
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: 60/011,787
FILING DATE: 16-FEB-1996
APPLICATION NUMBER: 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/386,844
FILING DATE: 10-FEB-1995

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-114-999
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:
LENGTH: 1481 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-924-417-60

Query Match 68.4%; Score 39; DB 10; Length 1481;

Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPPOL 9
Db 8 RWPPPLL 14

RESULT 9

US-09-864-761-42815
Sequence 42815, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/006666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42815
LENGTH: 71

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: MAP TO AC009946.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
OTHER INFORMATION: EST HUMAN HIT: AA873561.1, EVALUE 2.00e-20
OTHER INFORMATION: SWISSPROT HIT: P10069, EVALUE 1.90e+00
US-09-864-761-42815

Query Match 66.7%; Score 38; DB 10; Length 71;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
Db 40 PAWPPP 45

RESULT 10

US-09-764-898-188
; Sequence 188, Application US/09764898
; Patent No. US20020090673A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: P201

; CURRENT APPLICATION NUMBER: US/09/764, 898

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 311

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 188

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-898-188

Query Match 66.7%; Score 38; DB 10; Length 348;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
Db 142 PTWPPP 147

RESULT 11

US-09-764-853-568

; Sequence 568, Application US/09764853

; Patent No. US20020090672A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: P206

; CURRENT APPLICATION NUMBER: US/09/764, 853

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 939

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 568

; LENGTH: 482

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (188)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (194)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (215)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (281)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-853-568

Query Match 66.7%; Score 38; DB 10; Length 482;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPPP 7
| | | | |
Db 251 PTWPPP 256

RESULT 12

US-09-866-582-36

; Sequence 36, Application US/09866582

; Patent No. US20020127620A1

; GENERAL INFORMATION:

; APPLICANT: Witman, George B.

; APPLICANT: Pazour, Gregory J.

; APPLICANT: Rosenbaum, Joel L.

; APPLICANT: Cole, Douglas G.

; TITLE OF INVENTION: INTRAFAGELLAR TRANSPORT

; FILE REFERENCE: 07917-145001

; CURRENT APPLICATION NUMBER: US/09/866, 582

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 60/206, 923

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 36

; LENGTH: 1198

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-582-36

Query Match 66.7%; Score 38; DB 10; Length 1198;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPP 7
: | | | | |
Db 1140 IKRWPPP 1146

RESULT 13

US-10-001-215-5

; Sequence 5, Application US/10001215

; Patent No. US20020147323A1

; GENERAL INFORMATION:

; APPLICANT: Bandaru Rajasekhkar

; APPLICANT: Rosana Kapeller-Libermann

; TITLE OF INVENTION: 16224 and 69611, NOVEL HUMAN KINASES AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: MNI-209

; CURRENT APPLICATION NUMBER: US/10/001,215

; CURRENT FILING DATE: 2001-11-30

; PRIOR APPLICATION NUMBER: 60/250,917

; PRIOR FILING DATE: 2000-11-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 1241

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-001-215-5

Query Match 66.7%; Score 38; DB 12; Length 1241;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPRWPPP 7
: | | | | |
Db 1183 IKRWPPP 1189

RESULT 14

US-09-864-761-46140

; Sequence 46140, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46140
LENGTH: 39
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC023078.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EST HUMAN HIT: BE734169.1, EVALUE 9.00e-09
OTHER INFORMATION: SWISSPROT HIT: P31895, EVALUE 3.70e+00

US-09-864-761-46140
Query Match 64.9%; Score 37; DB 10; Length 39;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RWPPP 7
|||||

Db 22 RWPPP 26

RESULT 15
US-09-864-761-40064
Sequence 40064, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 40064
LENGTH: 52
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004408.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.8
OTHER INFORMATION: EST HUMAN HIT: AW952275.1, EVALUE 2.00e-12
OTHER INFORMATION: SWISSPROT HIT: P03204, EVALUE 2.70e+00

US-09-864-761-40064
Query Match 64.9%; Score 37; DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PRWPP 6
IIII
Db 42 PRWPP 46

Search completed: November 26, 2002, 19:44:26
Job time : 4.74194 secs

THIS PAGE BLANK (SPT0)

C	1	48.4	3.5	2000	9	US-09-938-842A-3267	Sequence 3267, App
C	2	48.4	3.5	2004	10	US-09-887-576-291	Sequence 211, App
C	3	48.4	3.5	2004	10	US-09-887-576-291	Sequence 211, App
C	4	45.6	3.3	257	10	US-09-860-352-14300	Sequence 297, App
C	5	44.6	3.2	1967	10	US-09-860-352-14300	Sequence 14300, App
C	6	43.4	3.1	2000	9	US-09-925-301-565	Sequence 565, App
C	7	42.8	3.1	239	10	US-09-938-842A-4451	Sequence 4451, App
C	8	42.6	3.1	324	10	US-09-960-352-11438	Sequence 11438, App
C	9	42.2	3.1	3238	10	US-09-960-352-11438	Sequence 10067, App
C	10	41.8	3.0	2136	9	US-09-874-528-5	Sequence 5, App
C	11	41.8	3.0	2136	9	US-09-992-698-302	Sequence 302, App
C	12	41.8	3.0	2136	10	US-09-989-722-302	Sequence 302, App
C	13	41.8	3.0	2136	10	US-09-989-723-302	Sequence 302, App
C	14	41.8	3.0	2136	10	US-09-989-279-302	Sequence 302, App
C	15	41.8	3.0	2136	10	US-09-989-727-302	Sequence 302, App
C	16	41.8	3.0	2136	10	US-09-989-731-302	Sequence 302, App
C	17	41.8	3.0	2136	10	US-09-989-732-302	Sequence 302, App
C	18	41.8	3.0	2136	10	US-09-991-073-302	Sequence 302, App
C	19	41.8	3.0	2136	10	US-09-990-442-302	Sequence 302, App
C	20	41.8	3.0	2136	10	US-09-991-163-302	Sequence 302, App

Qy	389	ACCTCTAC	396
Db	74	ATATCTAC	67

RESULT 5

US-09-925-301-565
; Sequence 565, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 565
; LENGTH: 1967
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-565

[illegible]

RESULT 6
 US-09-938-842A-4451
 : Sequence 4451, Application US/09938842A
 : Patent No. US20020160378A1
 : GENERAL INFORMATION:
 : APPLICANT: Harper, Jeff
 : APPLICANT: Kreps, Joel
 : APPLICANT: Wang, Xun
 : APPLICANT: Zhu, Tong
 : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 : TITLE OF INVENTION: SAME, AND METHODS OF USE
 : FILE REFERENCE: SCRIP1300-3
 : CURRENT APPLICATION NUMBER: US/09/938, 842A
 : CURRENT FILING DATE: 2001-08-24
 : PRIOR APPLICATION NUMBER: US 60/227, 866
 : PRIOR FILING DATE: 2000-08-24
 : PRIOR APPLICATION NUMBER: US 60/264, 647
 : PRIOR FILING DATE: 2001-01-16
 : PRIOR APPLICATION NUMBER: US 60/300, 111
 : PRIOR FILING DATE: 2001-06-22
 : NUMBER OF SEQ ID NOS: 5379
 : SEQ ID NO 4451
 : LENGTH: 2000
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 US-09-938-842A-4451

Query Match 3.1%; Score 43.4; DB 9; Length 2000;
Best Local Similarity 51.3%; Pred. No. 0.5;

	Matches	101:	Conservative	0:	Mismatches	96:	Indels	0:	Gaps	0:
Qy	88	AATAATTTAAACAGTACGCGGAAATCGACAGACAAATTTCTCCTCTCGTCACAAATATAGTGT	147							
Db	546	AAAAATATTTATAGTCATCAATGCTTATCGATTTTAAATATAGTTTAAAGATTATAATTT	605							
Qy	148	CAACACATTGGAAGTGAATTTTAAAGATGTTTAAATTTAAAGGATGGATTTCCAAAG	207							
Db	606	TGTGTTTTTGAATTTTAGATTTTGGGTTTTAGATTTTCCAAAAATAAAGATATATATAAA	665							
Qy	208	GA AAAAATAATAGGAAGGAAAGAAAAAACTGAACAGAAAAACGCAAAAGTATCATGTTTG	267							
Db	666	AAAAAATAGGAGAGAAAAAGTCACGTTATATTTGTTAAGAAAAAGCGAAAAACTCAATTTG	725							
Qy	268	GTCCTAAACCTTTTGCAA	284							
Db	726	AAAACTCAATTTTGCTA	742							

RESULT 7

```

RESULT /
US-09-960-352-11438/c
; Sequence 11438, Application US/09960352
; Patent NO. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagapan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES A
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11438
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 49-LIB3058-019-Q1-k1-E10
US-09-960-352-11438

```

Query Match	3.1%	Score 42.8	DB 10	Length 239
Best Local Similarity	52.2%	Pred. No. 0.25		
Matches	95	Conservative	0	Mismatches 87
				Indels 0
				Gaps 0
Qy	75	AATGAAGGTTGAAATAATTTAAACAGTACGGGAAATGCAGAGCAATTTTCTCCCTCTGGT	134	
Db	235	AATTTAAATAGATATTTATTACTTTTAAGCCAAAGTGAAGCCCTATTAGTGTTTTTT	176	
Qy	135	GACATATAGTGTCCAACACTTTGGAAAGTGATTTTTPAAGAATGTTTATTTTAAATTTAAAGG	194	
Db	175	TATATAGAATATTTCTACATTTTAAATTTCTTTTAAAAAATAAAAAAAAAAAAAA	116	
Qy	195	ATGGATTTCCCAAGGAAATAAATAATAGGAAAGAAAGAAACCTGAACAGAAACGCAA	254	
Db	115	AAACAA	56	
Qy	255	AA	256	
Db	55	AA	54	

RESULTS 8

US-09-960-352-10067/c
 : Sequence 10067, Application US/09960352
 : Patent No. US20020137139A1
 : GENERAL INFORMATION:
 : APPLICANT: Warren, Wesley C.
 : APPLICANT: Tao, Nengbing
 : APPLICANT: Byatt, John C.
 : APPLICANT: Mathialagan, Nagappan
 : TITLE OF INVENTION: NUCLEIC ACID AND O
 : TITLE OF INVENTION: MUSCLE AND FAT O

FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 10067
LENGTH: 364
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 43-LIB34-031-Q1-E1-C4
US-09-960-352-10067

Query Match 3.1%; Score 42.6; DB 10; Length 364;
Best Local Similarity 49.3%; Pred. No. 0.34;
Matches 111; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 35 GAGCAGGTACAAAAGTGTGTGCAATCTACITGGATGAAGGTTGAAATATT 94
DB 312 GACCAGGGGATAATGGCTGCTGGATTCAGATGATCTGTAGGAATCTTTGGCCATCAA 253
QY 95 TAAACAGTACGGGAAATGACAGCAATTTCTCTCTGTCGATGACATATAGTGTCCACAC 154
DB 252 CATCCAGTTTCTGAATAACCACTCTCTGTGTTCTCTGTGTAAGTACACAGTGTGCCCAT 193
QY 155 TTGGAAGTGAATTTAAGAAATTTATTTAAATTAAGAGTGGATTTCCAGGAAAAA 214
DB 192 GCGGGTGAATGGTGTCTAGTGTAGTTGTATTAACACTTTTGTGTTTCTCAAAAAA 133
QY 215 RATAAGGAAAGGAAAGAAAAAAGTGAACAGCAAAACGCAAAAGTA 259
DB 132 AAAAAAGAAAAATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 88

RESULT 9
US-09-874-628-5
Sequence 5, Application US/09874628
Patent No. US20020137133A1
GENERAL INFORMATION:
APPLICANT: WOZNEY, John
CELESTE, Anthony J.
THIES, R. Scott
YAMAJI, No. US20020137133A1
TITLE OF INVENTION: RECEPTOR PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute Inc.- Legal Affairs
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/874,628
FILING DATE: 05-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,934
FILING DATE: 17-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5203
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 876 1170
TELEFAX: 617 876 5851
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3238 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: CFK1-10a
FEATURE:
NAME/KEY: CDS
LOCATION: 474..2000
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-874-628-5
Query Match 3.1%; Score 42.2; DB 10; Length 3238;
Best Local Similarity 50.8%; Pred. No. 1.2;
Matches 101; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 51 GTGAGTGTGTGCAATCTTACTTGGAAATGAAAGGTTGAAATAATTTAAACAGTACGGGAAA 110
DB 3021 GGGGGGNAATGCATCTCTTCGGAACATATCCATTACATGCTTTAACTCTGCCAGAAA 3080
QY 111 TGCAGAGCAATTTCTCTCTCTGGTGACAATATAGTGTCCACACTTGGAAAGTATTTTTA 170
DB 3081 AAAAAATAACTATTTTGTGTTTAAATCTACTTTTGTATTTAGTAGTATTTCGTATAAATTA 3140
QY 171 AGAATGTTTATTTAAATTAAGGATGGATTTCCAGGAAAAAATAAGAAAAAGGAAA 230
DB 3141 AATAAATGTTTTCAGTCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAA 3200
QY 231 GAAAAAACTGAACAGAAAA 249
DB 3201 AAAAAAATAAAAAA 3219

RESULT 10
US-09-992-598-302
Sequence 302, Application US/09992598
Patent No. US20020160384A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311

1	PRIOR APPLICATION NUMBER: 60/089532
2	PRIOR FILING DATE: 1998-06-17
3	PRIOR APPLICATION NUMBER: 60/089538
4	PRIOR FILING DATE: 1998-06-17
5	PRIOR APPLICATION NUMBER: 60/089598
6	PRIOR FILING DATE: 1998-06-17
7	PRIOR APPLICATION NUMBER: 60/089599
8	PRIOR FILING DATE: 1998-06-17
9	PRIOR APPLICATION NUMBER: 60/089600
10	PRIOR FILING DATE: 1998-06-17
11	PRIOR APPLICATION NUMBER: 60/089653
12	PRIOR FILING DATE: 1998-06-17
13	PRIOR APPLICATION NUMBER: 60/089801
14	PRIOR FILING DATE: 1998-06-18
15	PRIOR APPLICATION NUMBER: 60/089907
16	PRIOR FILING DATE: 1998-06-18
17	PRIOR APPLICATION NUMBER: 60/089908
18	PRIOR FILING DATE: 1998-06-18
19	PRIOR APPLICATION NUMBER: 60/089947
20	PRIOR FILING DATE: 1998-06-19
21	PRIOR APPLICATION NUMBER: 60/089948
22	PRIOR FILING DATE: 1998-06-19
23	PRIOR APPLICATION NUMBER: 60/089952
24	PRIOR FILING DATE: 1998-06-19
25	PRIOR APPLICATION NUMBER: 60/090246
26	PRIOR FILING DATE: 1998-06-22
27	PRIOR APPLICATION NUMBER: 60/090252
28	PRIOR FILING DATE: 1998-06-22
29	PRIOR APPLICATION NUMBER: 60/090254
30	PRIOR FILING DATE: 1998-06-22
31	PRIOR APPLICATION NUMBER: 60/090349
32	PRIOR FILING DATE: 1998-06-23
33	PRIOR APPLICATION NUMBER: 60/090355
34	PRIOR FILING DATE: 1998-06-23
35	PRIOR APPLICATION NUMBER: 60/090429
36	PRIOR FILING DATE: 1998-06-24
37	PRIOR APPLICATION NUMBER: 60/090431
38	PRIOR FILING DATE: 1998-06-24
39	PRIOR APPLICATION NUMBER: 60/090435
40	PRIOR FILING DATE: 1998-06-24
41	PRIOR APPLICATION NUMBER: 60/090444
42	PRIOR FILING DATE: 1998-06-24
43	PRIOR APPLICATION NUMBER: 60/090445
44	PRIOR FILING DATE: 1998-06-24
45	PRIOR APPLICATION NUMBER: 60/090472
46	PRIOR FILING DATE: 1998-06-24
47	PRIOR APPLICATION NUMBER: 60/090542
48	PRIOR FILING DATE: 1998-06-24
49	PRIOR APPLICATION NUMBER: 60/090557
50	PRIOR FILING DATE: 1998-06-24
51	PRIOR APPLICATION NUMBER: 60/090690
52	PRIOR FILING DATE: 1998-06-25
53	PRIOR APPLICATION NUMBER: 60/090696
54	PRIOR FILING DATE: 1998-06-25
55	PRIOR APPLICATION NUMBER: 60/090678
56	PRIOR FILING DATE: 1998-06-25
57	PRIOR APPLICATION NUMBER: 60/090694
58	PRIOR FILING DATE: 1998-06-25
59	PRIOR APPLICATION NUMBER: 60/090695
60	PRIOR FILING DATE: 1998-06-25
61	PRIOR APPLICATION NUMBER: 60/090696
62	PRIOR FILING DATE: 1998-06-25
63	PRIOR APPLICATION NUMBER: 60/090682
64	PRIOR FILING DATE: 1998-06-26
65	PRIOR APPLICATION NUMBER: 60/090863
66	PRIOR FILING DATE: 1998-06-26
67	PRIOR APPLICATION NUMBER: 60/091360
68	PRIOR FILING DATE: 1998-07-01
69	PRIOR APPLICATION NUMBER: 60/091478
70	PRIOR FILING DATE: 1998-07-01

Page 6

[illegible]

[illegible]

```

RESULT 12
US-09-989-723-302
/ Sequence 302, Application US/09989723
/ Patent No. US2002072092A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Bostein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Kljavin, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Secreted and Tra
/ TITLE OF INVENTION: Acids Encoding
/ FILE REFERENCE: P2730PlC62
/ CURRENT APPLICATION NUMBER: US/09/98
/ CURRENT FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/049787
/ PRIOR FILING DATE: 1997-06-16
/ PRIOR APPLICATION NUMBER: 60/62250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/065186

```

PRIOR APPLICATION NUMBER: 60/089514	PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653	PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801	PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907	PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908	PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947	PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948	PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952	PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246	PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252	PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254	PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349	PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355	PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090557	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090540	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542	PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090676	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696	PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862	PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863	PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360	PRIOR FILING DATE: 1998-06-26

```

; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      3.0%; Score 41.8; DB 10; Length 2136;
Best Local Similarity 56.0%; Pred. No. 1.3;
Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 116 AGCAATTTTCCTCTGCTGACATATAGTGTCCACACTTGGAGTGAATTTTAAAGAA 175
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1979 AACTGTATCATTTCTGCTGAGGCTGAGTGTCCCATCTCTTTAACTCAAGTGTGGA 2038

Qy 176 GTTTATTAAATTAAGAGTGGATTTCCCAAGGAAAAAATAAGGAAAGGAAAGAAA 235
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2039 TTTTGACTAATAAAGAAAGATTTCTTAAGAAAAAAGAAAAAAGAAAAAAGAAA 2098

Qy 236 AACTGACAGACAGACGCAAAA 256
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2099 AAAAAAAAAAAAAAAAAAAAAA 2119

RESULT 13
US-09-989-279-302
; Sequence 302, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
```

;
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090866
; PRIOR FILING DATE: 1998-06-26

;
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09
;
Query Match 3.0%; Score 41.8; DB 10; Length 2136;
Best Local Similarity 56.0%; Pred. No. 1.3; Mismatches 0; Indels 0; Gaps 0;
Matches 79; Conservative 0;
Qy 116 AGCAATTTCTCTCTGTCGACAAATATAGTGTCCCAACACTTGGAGTGTATTTTAAAGAT 175
Db 179 AACTGTATCATTTTCTGCTGAGGTGGAGTGTCCCATCTTTTATCAAGGTGATTGTGA 2038
Qy 176 GTTTATTTTAAATTAAGGATGGATTTTCCAGGAAAAAATAAGGAAAGGAAAGAAAA 235
Db 2039 TTTTGACTAATAAAAAACAATTTGTAAAAAATAAAAAAATAAAAAAATAAAAAA 2098
Qy 236 AACTGACAGAAACGCAAAA 256
Db 2099 AAAAAAATAAAAAAATAAAAAA 2119
;
RESULT 14
US-09-989-727-302
; Sequence 302, Application US/09989727
; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerriksen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PLC65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

THIS PAGE BLANK (U3PT0)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 10:02:16 ; Search time 2767 Seconds
(without alignments)
14535.626 Million cell updates/sec

Title: US-09-674-593-1
Perfect score: 1382
Sequence: 1 cattaagctaacagcataaa.....taaaagcaattaagtctctgg 1382

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	1382	100.0	1382	9	AF181722	AF181722 Homo sapi	
c	2	1382	100.0	4377	9	AF181720	AF181720 Homo sapi
	3	1382	100.0	120029	2	HSJ282H10	AL132672 Homo sapi
4	1380.4	99.9	152966	9	AL359713	AL359713 Human DNA	
c	5	682.4	49.4	1287	9	BC014954	BC014954 Homo sapi
c	6	595.4	43.1	2167	9	AF181721	AF181721 Homo sapi
c	7	409.8	29.7	2003	9	AK027036	AK027036 Homo sapi
c	8	293.4	21.2	6548	9	AB032980	AB032980 Homo sapi
9	252.8	18.3	199957	10	AL589735	AL589735 Mouse DNA	
10	86.2	6.2	169103	2	AC125638	AC125638 Rattus no	
c	11	85.4	6.2	139357	9	AC010907	AC010907 Homo sapi
c	12	75.8	5.5	115756	9	HSJ262L5	AL049795 Human DNA
c	13	75.8	5.5	322840	2	HSJ2675E8	AL121991 Homo sapi
c	14	68	4.9	2422	6	AX056382	AX056382 Sequence
c	15	67.8	4.9	3295	3	AB047038	AB047038 Halocynthia
16	66.2	4.8	141790	10	AL671759	AL671759 Mouse DNA	
17	61.2	4.4	125020	9	AF429315	AF429315 Homo sapi	
c	18	60.4	4.4	125020	9	AF429315	AF429315 Homo sapi
c	19	55.4	4.0	77096	2	AC116030	AC116030 Dictyoste
c	20	55.2	4.0	27057	1	AF169302	AF169302 Burkholder
21	53.6	3.9	12965	1	AE002048	AE002048 Deinococcus	
c	22	53.4	3.9	104014	2	AC116921	AC116921 Dictyoste
c	23	52.8	3.8	47791	2	AC115593	AC115593 Dictyoste
c	24	52.8	3.8	47791	2	AC116305	AC116305 Dictyoste
c	25	51.6	3.7	11735	6	AX251838	AX251838 Sequence
c	26	51.6	3.7	11735	6	AX277870	AX277870 Sequence
c	27	51.6	3.7	11735	6	AX323553	AX323553 Sequence
c	28	51.4	3.7	156533	2	AC117070	AC117070 Dictyoste
c	29	51.2	3.7	182054	10	AL365314	AL365314 Mouse DNA
30	51.2	3.7	184754	2	AC022675	AC022675 Mus musculus	
31	50.8	3.7	150613	8	AB001684	AB001684 Chlorella	
32	50.8	3.7	175699	2	AC129136	AC129136 Rattus norvegicus	
33	50.2	3.6	132000	2	AC116976	AC116976 Dictyoste	
34	50	3.6	124820	2	AC117073	AC117073 Dictyoste	
35	50	3.6	199450	9	AC005674	AC005674 Homo sapi	
c	36	49.4	3.6	20043	2	AC116986	AC116986 Dictyoste
c	37	49.4	3.6	106144	2	AC116425	AC116425 Dictyoste
38	49.2	3.6	3671	3	AF481922	AF481922 Dictyoste	
c	39	49.2	3.6	33156	2	AC116028	AC116028 Dictyoste
40	49.2	3.6	100998	9	HS343K2	AL031659 Human DNA	
41	49.2	3.6	195896	2	AL805976	AL805976 Mus musculus	
42	49	3.5	2872	3	DDIUBIRPB	M19492 Slime mold	
43	49	3.5	93491	2	AC116967	AC116967 Dictyoste	
c	44	49	3.5	94754	2	AC103188	AC103188 Rattus norvegicus
45	49	3.5	173124	2	AC099264	AC099264 Rattus norvegicus	

ALIGNMENTS

RESULT 1	AF181722	1382 bp	mRNA	linear	PRI 10-JAN-2000
LOCUS	Homo sapiens RU2AS (RU2) mRNA, complete cds.				
DEFINITION	AF181722				
ACCESSION	AF181722.1	GI:6684531			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1382)				
AUTHORS	Van Den Eynde,B.J., Gaugler,B., Probst-Kepper,M., Michaux,L.,				
	Devuyt,O., Lorge,F., Weynants,P. and Boon,T.				
TITLE	A new antigen recognized by cytolytic T lymphocytes on a human				

kidney tumor results from reverse strand transcription
J. Exp. Med. 190 (12), 1793-1800 (1999)
MEDLINE 20069887
PUBMED 10601354
REFERENCE 2 (bases 1 to 1382)
AUTHORS Van den Eynde,B.J., Gaugler,B. and Pilotte,L.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-1999) Ludwig Institute for Cancer Research,
Avenue Hippocrate, 74, Brussels 1200, Belgium
FEATURES
 source Location/Qualifiers
 1..1382
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p22.1"
 1..1382
 /gene="RU2"
 738..992
 /gene="RU2"
 /codon_start=1
 /product="RU2AS"
 /protein_id="AAF23613.1"
 /db_xref="GI:6684532"
 /translation="MDDDAAPRVGVVAVHKHALHDGLRQVAGGAAAHLPWP
 QLAASRREAPLSPRPRTQAGSPPTNEKLTPQVKEK"
BASE COUNT 355 a 373 c 344 g 310 t
ORIGIN

Query Match 100.0%; Score 1382; DB 9; Length 1382;
Best Local Similarity 100.0%; Pred. No. 1.1e-306;
Matches 1382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATTATGCTAACAGCATAAACATGACAGGGGTGGGAGGAGGTGCACAAAAGTGAGTGTG 60
DB 1 CATTATGCTAACAGCATAAACATGACAGGGGTGGGAGGAGGTGCACAAAAGTGAGTGTG 60

QY 61 TCAATTCTACTTGGAAATGAAGTTGAATAATTTAAACAGTACGCGGAAATGCAGACAA 120
DB 61 TCAATTCTACTTGGAAATGAAGTTGAATAATTTAAACAGTACGCGGAAATGCAGACAA 120

QY 121 TTTTCTCTCTGGTGACAAATAGTGTCCACACATCTGGAAGTGATTTTAAAGAATGTTTA 180
DB 121 TTTTCTCTCTGGTGACAAATAGTGTCCACACATCTGGAAGTGATTTTAAAGAATGTTTA 180

QY 181 TTTAAATTAAGAGTGGATTTCACGAAAGAAATTAAGGAAAGGAAAGAAAGAACTG 240
DB 181 TTTAAATTAAGAGTGGATTTCACGAAAGAAATTAAGGAAAGGAAAGAAAGAACTG 240

QY 241 AACAGAAACGCAAAAGTATCAGTTTGGTCACTAACCTTTGCAAGGATACCTTTTATTT 300
DB 241 AACAGAAACGCAAAAGTATCAGTTTGGTCACTAACCTTTGCAAGGATACCTTTTATTT 300

QY 301 TCTTTAAGATTCCTGTTTATACACAGATTTTAAAGTTTACTCTACTGTCGCCCAAG 360
DB 301 TCTTTAAGATTCCTGTTTATACACAGATTTTAAAGTTTACTCTACTGTCGCCCAAG 360

QY 361 TGAATATCTCTCCAGTCACAGTCTCACTCTCACTCCCTCCCTCACTCACTCACTCACT 420
DB 361 TGAATATCTCTCCAGTCACAGTCTCACTCTCACTCCCTCCCTCACTCACTCACTCACT 420

QY 421 AGGGGCATCAATCACACCGGAGAGTCAACAGCCCTCAACACCTGAGGTGTGGGGGGTAG 480
DB 421 AGGGGCATCAATCACACCGGAGAGTCAACAGCCCTCAACACCTGAGGTGTGGGGGGTAG 480

QY 481 GGATCTGATTTCTTCATATCAACCCACACTATAGGACCTAAATGGTGGGGGGTAG 540
DB 481 GGATCTGATTTCTTCATATCAACCCACACTATAGGACCTAAATGGTGGGGGGTAG 540

QY 541 GGGAGACCGACTCACTTGAGTTCTTGAAGGCTTCCCTCCCTCCAGCCACGTAATTTGCC 600
DB 541 GGGAGACCGACTCACTTGAGTTCTTGAAGGCTTCCCTCCCTCCAGCCACGTAATTTGCC 600

QY 601 CCGCTCTGGATCTGCTAGCTTCCGGATTCGGTGCCAGTCCGCGGGGTGTAGATGTTC 660

|||||
601 CCGCTCTGGATCTGCTAGCTTCCGGATTCGGTGCCAGTCCGCGGGGTGTAGATGTTC 660

QY 661 CTGACGGCCCAAAAGGGTGCCTGGAACGCGCGGGTCACTCTCTCAGGAAGACTTCGAAG 720
DB 661 CTGACGGCCCAAAAGGGTGCCTGGAACGCGCGGGTCACTCTCTCAGGAAGACTTCGAAG 720

QY 721 CTGGACACTTCTTCTCATATGATGACGACGCGCGCGCGCGTGAAGGGTCCCGGTG 780
DB 721 CTGGACACTTCTTCTCATATGATGACGACGCGCGCGCGCGTGAAGGGTCCCGGTG 780

QY 781 CGGTACACAAGCACGCTCTTACGACGGGTGAGACAGGTGCTGACCTGCGGTGTG 840
DB 781 CGGTACACAAGCACGCTCTTACGACGGGTGAGACAGGTGCTGACCTGCGGTGTG 840

QY 841 CCGCTCATCTTCCCGCTGGCGCGCGCTCAGCTCGCTGCTTCGGCTCGGGAGGACCTC 900
DB 841 CCGCTCATCTTCCCGCTGGCGCGCGCTCAGCTCGCTGCTTCGGCTCGGGAGGACCTC 900

QY 901 CGCTGTCCAGCGGCTCACCGCACCCAGGGCGGGGATCGCTCTCTGAAACGAACGAGA 960
DB 901 CGCTGTCCAGCGGCTCACCGCACCCAGGGCGGGGATCGCTCTCTGAAACGAACGAGA 960

QY 961 AACTGACGAATCCACAGGTGAAAGAGAAAGTAACGGCCGTGCGCTAGGCGTCCACCCAGA 1020
DB 961 AACTGACGAATCCACAGGTGAAAGAGAAAGTAACGGCCGTGCGCTAGGCGTCCACCCAGA 1020

QY 1021 GGAGACACTAGAGCTTTCAGGACTCGGAGTAGAGCTCAAGTTTTCACCGTGGCGTGC 1080
DB 1021 GGAGACACTAGAGCTTTCAGGACTCGGAGTAGAGCTCAAGTTTTCACCGTGGCGTGC 1080

QY 1081 ACAGCAATCAGGACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 ACAGCAATCAGGACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140

QY 1141 CAAGTGACAGCTTCTGAGCAGGAGCGGAAACGCGGGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 CAAGTGACAGCTTCTGAGCAGGAGCGGAAACGCGGGGCGGCGGCGGCGGCGGCGGCGG 1200

QY 1201 TAGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 TAGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260

QY 1261 CAATTTCTTACATATATCTGCTACCTTTCCAAACAGCCTAATTTTCTTTCTCTC 1320
DB 1261 CAATTTCTTACATATATCTGCTACCTTTCCAAACAGCCTAATTTTCTTTCTCTC 1320

QY 1321 TTTCTGACCTTTACCCCTCAATCTCCTGCTTCTCCCAAAATTAAGCAATTAAGTTCT 1380
DB 1321 TTTCTGACCTTTACCCCTCAATCTCCTGCTTCTCCCAAAATTAAGCAATTAAGTTCT 1380

QY 1381 GG 1382
DB 1381 GG 1382

RESULT 2
AF181720/c 4377 bp DNA linear PRI 10-JAN-2000
LOCUS Homo sapiens RU2AS (RU2) gene, complete cds; and RU2S (RU2) gene,
DEFINITION partial cds.
ACCESSION AF181720
VERSION AF181720.1 GI:6684526
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4377)
AUTHORS Van Den Eynde,B.J., Gaugler,B., Probst-Kepper,M., Michaux,L.,
Devuyt,O., Lorge,F., Weynants,P. and Boon,F.
TITLE A new antigen recognized by cytolytic T lymphocytes on a human
kidney tumor results from reverse strand transcription

J. Exp. Med. 190 (12), 1793-1800 (1999)
20069887
MEDLINE
10601354
PUBMED
2 (bases 1 to 4377)
REFERENCE
Van den Eynde,B.J., Gaugler,B. and Pilotte,L.
AUTHORS
Direct Submission
TITLE
Submitted (30-AUG-1999) Ludwig Institute for Cancer Research,
JOURNAL
Avenue Hippocrate, 74, Brussels 1200, Belgium
Location/Qualifiers
FEATURES
1..4377
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22.1"
/1754..>2348
gene
/gene="R2"
1754..>2348
mRNA
/gene="R2"
/product="R2"
complement(<1912..>2166)
mRNA
/product="R2AS"
complement(1912..2166)
CDS
/codon_start=1
/product="R2AS"
/protein_id="AAF23611.1"
/db_xref="GI:6684528"
/translation="MDDDAARVEGVVAVHKHALHDGLRQVAGPGAAAHLPWPPP
QLAASREAPPLSQRPHTRGAGSPPETNEKLTPQYKER"
2056..>2348
CDS
/gene="R2"
/codon_start=1
/product="R2"
/protein_id="AAF23610.1"
/db_xref="GI:6684527"
/translation="MSGSSARSHLSQPVKSVLYRNGDPFFYAGRVRVIEHKVVSF
EVLKEVIGVQAPFGAVRNIYPTRGHRIKKLDQIQSGGNYVAGGGEAPKKL"
BASE COUNT 1296 a 858 g 828 t 1380 c 15 others
ORIGIN
Query Match 100.0%; Score 1382; DB 9; Length 4377;
Best Local Similarity 100.0%; Pred. No. 1.4e-306;
Matches 1382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy - 1 CATTATGCTAACAGCATAAACATGCGAGGGGTGGAGCAGGGTGCACAAAAGTGAAGTGTG 60
Db 2903 CATTATGCTAACAGCATAAACATGCGAGGGGTGGAGCAGGGTGCACAAAAGTGAAGTGTG 2844
Oy 61 TCAATTCCTACTGGAATGAAAGGTTGAAATATTTAAACAGTACGGGAAATGCAGAGCAA 120
Db 2843 TCAATTCCTACTGGAATGAAAGGTTGAAATATTTAAACAGTACGGGAAATGCAGAGCAA 2784
Oy 121 TTTTCTCTCTGTCACAAATATAGTCCAACTTGGAACTTGAAGTATTTTAAAGATGTTTA 180
Db 2783 TTTTCTCTCTGTCACAAATATAGTCCAACTTGGAACTTGAAGTATTTTAAAGATGTTTA 2724
Oy 181 TTTAAATTTAAAGGATGGATTCCAAAGGAAAGAAATTAAGGAAAGGAAAGAAAGAACTG 240
Db 2723 TTTAAATTTAAAGGATGGATTCCAAAGGAAAGAAATTAAGGAAAGGAAAGAAAGAACTG 2664
Oy 241 AACAGAAACCCAAAGTATCAGTTTGTGTCACCTAACCTTTGCAAGGATACCTTTTATTT 300
Db 2663 AACAGAAACCCAAAGTATCAGTTTGTGTCACCTAACCTTTGCAAGGATACCTTTTATTT 2604
Oy 301 TCTTTAAGATCCCTGTGTTGTTATACACAGATTTTAAAGTATCTACTGCTGACCCCAAG 360
Db 2603 TCTTTAAGATCCCTGTGTTGTTATACACAGATTTTAAAGTATCTACTGCTGACCCCAAG 2544
Oy 361 TGAATTCCTCTCTCAGTCACAGTGTCAACCTCTACCCCACTGCAACAGAGAGTGTG 420
Db 2543 TGAATTCCTCTCTCAGTCACAGTGTCAACCTCTACCCCACTGCAACAGAGAGTGTG 2484
Oy 421 AGGGGCATCAATCAACCCGAGAGTACAGCCCTCAACCACTCAGAGTGTGGGGGGTAG 480
|||||

Db 2483 AGGGGCATCAATCAACCCGAGAGTACAGGCCCTCAACCACTGAJGTGTGGGGGGTAG 2424
Oy 481 GGATCTGATTTCTTCATATCAACCCCACTATAGGCACCTAAATGGGTGGGGGGTGG 540
|||||
Db 2423 GGATCTGATTTCTTCATATCAACCCCACTATAGGCACCTAAATGGGTGGGGGGTGG 2364
|||||
Oy 541 GGGAGACCGACTCACTTTCAGTTTCTTGAAGGCTTCTGGCCCTCCACACGCTAAATGGCC 600
|||||
Db 2363 GGGAGACCGACTCACTTTCAGTTTCTTGAAGGCTTCTGGCCCTCCACACGCTAAATGGCC 2304
|||||
Oy 601 CCGCTCTGGATCTGGTCTAGCTTCGGGATTCGGTGGCCAGTCCCGGGGTGTAGATGTC 660
|||||
Db 2303 CCGCTCTGGATCTGGTCTAGCTTCGGGATTCGGTGGCCAGTCCCGGGGTGTAGATGTC 2244
|||||
Oy 661 CTGAGCGCCCAAGGGGTGCGTGAACCGCGCGGTCACTCTCTTCAGGAAGACTTCGAAG 720
|||||
Db 2243 CTGAGCGCCCAAGGGGTGCGTGAACCGCGCGGTCACTCTCTTCAGGAAGACTTCGAAG 2184
|||||
Oy 721 CTGAGACCTTTCTTCATGATGACGACGCGGGCCCGCGGTAGAAGGGTTCGCCGTG 780
|||||
Db 2183 CTGAGACCTTTCTTCATGATGACGACGCGCGCGGTAGAAGGGTTCGCCGTG 2124
|||||
Oy 781 CCGTACACAAGACGCTCTTCACGACGGGTGAGACAGGTGGCTGGACTGGCGTCTG 840
|||||
Db 2123 CCGTACACAAGACGCTCTTCACGACGGGTGAGACAGGTGGCTGGACTGGCGTCTG 2064
|||||
Oy 841 CCGCTCATCTTCCCGCTGGCGCGCTCAGCTCGCTTCGCTCGGGAGGACCTC 900
|||||
Db 2063 CCGCTCATCTTCCCGCTGGCGCGCTCAGCTCGCTTCGCTTCGGGTTCGGAGGACCTC 2004
|||||
Oy 901 CGCTGTCCAGCGGCTCACCGACCCAGCGGGGATCGCTCTCTGAAAGAAACGAGA 960
|||||
Db 2003 CGCTGTCCAGCGGCTCACCGACCCAGCGGGGATCGCTCTCTGAAAGAAACGAGA 1944
|||||
Oy 961 AACTGACCAATCCACAGGTGAAAGAGTAAAGCGCCGTGGCCCTAGCGTCCACCCAGA 1020
|||||
Db 1943 AACTGACCAATCCACAGGTGAAAGAGTAAAGCGCCGTGGCCCTAGCGTCCACCCAGA 1884
|||||
Oy 1021 GGAGACACTAGGAGCTTCAGGACTCGCAGTAGAGGCTCAAGTTTTCACCTGGCGTGC 1080
|||||
Db 1883 GGAGACACTAGGAGCTTCAGGACTCGCAGTAGAGGCTCAAGTTTTCACCTGGCGTGC 1824
|||||
Oy 1081 ACAGCCAAATCAGGACCCGAGTGGCGCACCAACAGGTTCACCTGTACGGGAGAAAT 1140
|||||
Db 1823 ACAGCCAAATCAGGACCCGAGTGGCGCACCAACAGGTTCACCTGTACGGGAGAAAT 1764
|||||
Oy 1141 CAAGGTGACAGCTTCTCAGCAGGAGCGGAAACCGCGGGGCTTCAACAGGACGCC 1200
|||||
Db 1763 CAAGGTGACAGCTTCTCAGCAGGAGCGGAAACCGCGGGGCTTCAACAGGACGCC 1704
|||||
Oy 1201 TAGTGAGGCGAGGAGAGGAGGAGGACACACACACACACACAAATATGCTCAAAACC 1260
|||||
Db 1703 TAGTGAGGCGAGGAGAGGAGGAGGAGGACACACACACACACAAATATGCTCAAAACC 1644
|||||
Oy 1261 CAATTTCTTACATCATATCTGTGTACCTTTTCCAAACAGCCTAAATTTTCTTTCTCTC 1320
|||||
Db 1643 CAATTTCTTACATCATATCTGTGTACCTTTTCCAAACAGCCTAAATTTTCTTTCTCTC 1584
|||||
Oy 1321 TTCTTGACCTTTACCCCTCAATCTCTGTCTCTCCCAAAATTAAGCAATTAAGTTCT 1380
|||||
Db 1583 TTCTTGACCTTTACCCCTCAATCTCTGTCTCTCTCCCAAAATTAAGCAATTAAGTTCT 1524
|||||
Oy 1381 GG 1382
||
Db 1523 GG 1522

RESULT 3
HSJ282H10
LOCUS
DEFINITION Homo sapiens chromosome 6 clone RPI-282H10 map p22.1-22.3, ***
SEQUENCING IN PROGRESS ***, 5 unordered pieces.
ACCESSION AL132672

All32672.14 GI:14348905
HTG: HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 120029)
JOURNAL Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 12, 2001 this sequence version replaced gi:12331282.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj282H10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M77815; 0% of reads
Chemistry: plasmid; L08752; 99% of reads
Chemistry: Dye-terminator ABI; 1% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Consensus quality: 118224 bases at least Q40
Consensus quality: 118449 bases at least Q30
Consensus quality: 118634 bases at least Q20
Insert size: 119629; sum-of-contigs
Insert size: 117431; 9.3% error; agarose-fp
Quality coverage: 18.41x in Q20 bases; sum-of-contigs Quality
coverage: 19.36x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1. 23429: contig of 23429 bp in length
* 23430 23529: gap of 100 bp
* 23530 63600: contig of 40071 bp in length
* 63601 63700: gap of 100 bp
* 63701 100099: contig of 36399 bp in length
* 100100 100199: gap of 100 bp
* 100200 102366: contig of 2167 bp in length
* 102367 102466: gap of 100 bp
* 102467 120029: contig of 17563 bp in length.
*
* Location/Qualifiers
*
* 1. 120029
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="6"
* /map="p22.1-22.3"
* /clone="RP1-282H10"
* /clone_lib="RPC1-1"
*
* misc_feature
* 1. 23429
* /note="assembly_fragment:05837"
* clone_end:SP6
* vector_side:left
* 23530..63600
* /note="assembly_fragment:00423"
* fragment_chain:1
* 63701..100099
* /note="assembly_fragment:02685"
* fragment_chain:1
* 100200..102366
* /note="assembly_fragment:01789"
* 102467..120029
* /note="assembly_fragment:02341"

clone_end:T7
vector_side:right
BASE COUNT 36483 a 24840 c 24637 g 33666 t 403 others
ORIGIN
Query Match 100.0%; Score 1382; DB 2; Length 120029;
Best Local Similarity 100.0%; Pred. No. 2.8e-306;
Matches 1382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATTATGCTAACAGCATAAACATGCAGGGGTGGAGCGGTGCACAAAAGTGTGTTG 60
|||||
Db 21048 CATTATGCTAACAGCATAAACATGCAGGGGTGGAGCGGTGCACAAAAGTGTGTTG 21107
QY 61 TCAATCTTCTTGAAGGTTGAATAATTTAAACAGTACGGGAAATGCAGACAA 120
|||||
Db 21108 TCAATCTTCTTGAAGGTTGAATAATTTAAACAGTACGGGAAATGCAGACAA 21167
QY 121 TTTTCTCTCTGTGACAATATAGTGTCCAACTTGTGAAGTGTATTTTAAAGATGTTTA 180
|||||
Db 21168 TTTTCTCTCTGTGACAATATAGTGTCCAACTTGTGAAGTGTATTTTAAAGATGTTTA 21227
QY 181 TTTAAATTTAAAGGATGGATTTTCCAGGAAATAAGGAAAGGAAAGGAAAGGAAAGT 240
|||||
Db 21228 TTTAAATTTAAAGGATGGATTTTCCAGGAAATAAGGAAAGGAAAGGAAAGGAAAGT 21287
QY 241 AACGAAAGCGAAAGATATCAGTTTGGTCACTAACCTTTGCAAGGATACCTTTTATTT 300
|||||
Db 21288 AACGAAAGCGAAAGATATCAGTTTGGTCACTAACCTTTGCAAGGATACCTTTTATTT 21347
QY 301 TCTTTAAGATTCCTGTTGTTTATACACAGATTTTAAAGTTTACTCTCTGTCACCAAG 360
|||||
Db 21348 TCTTTAAGATTCCTGTTGTTTATACACAGATTTTAAAGTTTACTCTCTGTCACCAAG 21407
QY 361 TGAATTCCTTCTCCAGTCACAGTGTCAACCTTACCCCCCACTGCAACGAGTTTG 420
|||||
Db 21408 TGAATTCCTTCTCCAGTCACAGTGTCAACCTTACCCCCCACTGCAACGAGTTTG 21467
QY 421 AGGGCATCAATCACCGGAGAGTCACAGCCCTCAACCACTGAGTGTGGGGGGTAG 480
|||||
Db 21468 AGGGCATCAATCACCGGAGAGTCACAGCCCTCAACCACTGAGTGTGGGGGGTAG 21527
QY 481 GGATCTGATTTCTTCAATATCAACCCACATATAGGCGACCTAAATGGGTGGGGGTGG 540
|||||
Db 21528 GGATCTGATTTCTTCAATATCAACCCACATATAGGCGACCTAAATGGGTGGGGGTGG 21587
QY 541 GGGAGACCGACTCACTTCACTTCTTGAAGGCTTCCAGCCCTCCAGCCACGTAATTGCC 600
|||||
Db 21588 GGGAGACCGACTCACTTCACTTCTTGAAGGCTTCCAGCCCTCCAGCCACGTAATTGCC 21647
QY 601 CCGCTCTGGATCTGTGTCTAGCTTCCGGATTCCGGTCCGAGTCCGCGGGGTCTAGATGTC 660
|||||
Db 21648 CCGCTCTGGATCTGTGTCTAGCTTCCGGATTCCGGTCCGAGTCCGCGGGGTCTAGATGTC 21707
QY 661 CTGAGCGCCCAAGGGTGCCTTGAACGCCCGCGGTCACTCTCTCAGGAAGACTTCGAAG 720
|||||
Db 21708 CTGAGCGCCCAAGGGTGCCTTGAACGCCCGCGGTCACTCTCTCAGGAAGACTTCGAAG 21767
QY 721 CTGGACACCTTCTTCTCATGGATGACGACGGCGCGCGCTAGAGGGGTCCCGGTG 780
|||||
Db 21768 CTGGACACCTTCTTCTCATGGATGACGACGGCGCGCGCTAGAGGGGTCCCGGTG 21827
QY 781 CGGTACACAAGCAGCTCTTACAGCGGGCTGAGACAGTGGCTGACCTGGCGTGTG 840
|||||
Db 21828 CGGTACACAAGCAGCTCTTACAGCGGGCTGAGACAGTGGCTGACCTGGCGTGTG 21887
QY 841 CCGCTCATCTTCCCGCTGGCGCGCTCAGTGTGCTCTTCCGCTCGGGGACCTTC 900
|||||
Db 21888 CCGCTCATCTTCCCGCTGGCGCGCTCAGTGTGCTCTTCCGCTCGGGGACCTTC 21947
QY 901 CCGTCTCCAGCGGCTCACCGCACCCAGGCGGGATCGCTCTCTGAAACGACGAGA 960
|||||
Db 21948 CCGTCTCCAGCGGCTCACCGCACCCAGGCGGGATCGCTCTCTGAAACGACGAGA 22007

Qy 961 AACTGACGAATCCACAGGTGAAGAGAGAGTAACGCCCGCTGCGCTAGCGTCCACCCCGAGA 1020
 |||
 Db 22008 AACTGACGAATCCACAGGTGAAGAGAGAGTAACGCCCGCTGCGCTAGCGTCCACCCCGAGA 22067
 |||
 Qy 1021 GGAGACACTAGGAGCTTCAGGACTCGGAGTAGAGGCTCAAGTTTTCACCGTGGCGTGC 1080
 |||
 Db 22068 GGAGACACTAGGAGCTTCAGGACTCGGAGTAGAGGCTCAAGTTTTCACCGTGGCGTGC 22127
 |||
 Qy 1081 ACAGCCAATCAGGACCCGCGAGTGGCGGACACACACAGGTTTCACCTGCTACGGGCGAGAAT 1140
 |||
 Db 22128 ACAGCCAATCAGGACCCGCGAGTGGCGGACACACACAGGTTTCACCTGCTACGGGCGAGAAT 22187
 |||
 Qy 1141 CAAGGTGACAGCTTCGTGAGCAGAGCGCGGAACGCGGGGCGCTTCAACAGGACGCC 1200
 |||
 Db 22188 CAAGGTGACAGCTTCGTGAGCAGAGCGCGGAACGCGGGGCGCTTCAACAGGACGCC 22247
 |||
 Qy 1201 TAGTGAGGCGAGGAGGAGGAGCGACACACACACACACACACAATATGGTCAACCC 1260
 |||
 Db 22248 TAGTGAGGCGAGGAGGAGGAGCGACACACACACACACACAATATGGTCAACCC 22307
 |||
 Qy 1261 CAATTTCTTACATCATCTCTGTCTACCTTTTCCAAACAGCGCTAATTTTCTTTCTCTC 1320
 |||
 Db 22308 CAATTTCTTACATCATCTCTGTCTACCTTTTCCAAACAGCGCTAATTTTCTTTCTCTC 22367
 |||
 Qy 1321 TTCTTGACCTTTACCCCTCAATCTCTGCTCTCTCCCAAAATTAAGCAATTAAGTTCCT 1380
 |||
 Db 22368 TTCTTGACCTTTACCCCTCAATCTCTGCTCTCTCCCAAAATTAAGCAATTAAGTTCCT 22427
 |||
 Qy 1381 GG 1382
 ||
 Db 22428 GG 22429

RESULT 4
 AL359713
 LOCUS Human DNA sequence from clone RP11-95P3 on chromosome 6, complete
 DEFINITION
 AL359713
 VERSION AL359713.25 GI:13938809
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 152566)
 Kimberley, A.
 Direct Submission
 Submitted (01-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On May 3, 2001 this sequence version replaced gi:13446455.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em:, EMBL; Sw:,
 SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6
 RP11-95P3 is from the library RPI-11.1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-95P3 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-95P3 is at 1 in this sequence. The
 true left end of clone RPI-73M23 is at 152467 in this sequence. The
 true right end of clone RP11-40E20 is at 17700 in this sequence.

FEATURES

source
 1..152966
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-95P3"
 /clone_lib="RPI-11.1"
 19..544
 /note="match: GSS: Em:AZ517849"
 24..379
 /note="match: GSS: Em:B63526"
 32..494
 /note="match: GSS: Em:AQ315706"
 70..589
 /note="match: GSS: Em:AQ285677"
 179..362
 /note="MIR repeat: matches 65..262 of consensus"
 521..582
 /note="L2 repeat: matches 2683..2741 of consensus"
 720..1022
 /note="AlusX repeat: matches 1..303 of consensus"
 1327..1448
 /note="AluJo/FRAM repeat: matches 181..298 of consensus"
 2264..2575
 /note="Alusq repeat: matches 1..311 of consensus"
 3014..3296
 /note="match: STS: Em:G27290"
 3014..3093
 /note="40 copies 2 mer ga 75% conserved"
 5024..5477
 /note="L1M1 repeat: matches 1012..1598 of consensus"
 5581..5859
 /note="Alusg repeat: matches 18..294 of consensus"
 5986..6359
 /note="L1MEC repeat: matches 2228..2259 of consensus"
 6367..6515
 /note="FLAM_C repeat: matches 1..143 of consensus"
 6971..7198
 /note="L2 repeat: matches 1512..1735 of consensus"
 8403..8838
 /note="LTR7 repeat: matches 1..445 of consensus"
 8799..9441
 /note="match: GSS: Em:B54691"
 complement(9393..9853)
 /note="match: GSS: Em:AQ285817"
 10311..10437
 /note="L2 repeat: matches 2612..2747 of consensus"
 10710..11022
 /note="AluJo repeat: matches 1..309 of consensus"
 12006..12156
 /note="L2 repeat: matches 2594..2748 of consensus"
 12182..13011
 /note="L2 repeat: matches 1374..2344 of consensus"
 13404..13631
 /note="L14 copies 2 mer at 55% conserved"
 13479..13630
 /note="38 copies 4 mer tata 60% conserved"
 13785..14820
 /note="L1ME repeat: matches 4766..5798 of consensus"
 14928..16008
 /note="L1PB3 repeat: matches 5106..6149 of consensus"
 16029..16110

```

/note="L1PB3 repeat: matches 5054. .5133 of consensus"
16428. .16829
repeat_region
/note="L1MC3 repeat: matches 6189. .6603 of consensus"
16868. .17150
repeat_region
/note="AluX repeat: matches 1. .285 of consensus"
17246. .17991
repeat_region
/note="L1MC4 repeat: matches 6608. .7343 of consensus"
17992. .18675
repeat_region
/note="MER67C repeat: matches 1. .710 of consensus"
18676. .18729
repeat_region
/note="L1MC4 repeat: matches 7343. .7396 of consensus"
18731. .19458
repeat_region
/note="MER4D repeat: matches 232. .973 of consensus"
complement(1929. .19542)
/note="match: GSS: Em:AQ342961"
complement(19031. .19542)
/note="match: GSS: Em:AQ053552"
complement(19065. .19495)
/note="match: GSS: Em:AQ545892"
complement(19130. .19546)
/note="match: GSS: Em:B94529"
19255. .19495
/note="match: GSS: Em:AQ373351 Em:AQ375803
match: STS: Em:G59266"
complement(19269. .19546)
/note="match: GSS: Em:AQ663911"
complement(19325. .19546)
/note="match: GSS: Em:AQ021494"
complement(19337. .19501)
/note="match: GSS: Em:AQ268095"
complement(19370. .19546)
/note="match: GSS: Em:AQ128461"
19566. .19739
/note="AluSg/x repeat: matches 126. .295 of consensus"
19740. .20100
repeat_region
/note="L1MD3 repeat: matches 7391. .7739 of consensus"
21121. .21429
/note="AluJo repeat: matches 1. .310 of consensus"
22589. .22728
/note="70 copies 2 mer aa 60% conserved"
22591. .22670
/note="20 copies 4 mer aaag 80% conserved"
22686. .22761
/note="19 copies 4 mer aagg 93% conserved"
23098. .23393
/note="AluJb repeat: matches 4. .298 of consensus"
23468. .23578
/note="L2 repeat: matches 2580. .2700 of consensus"
23980. .24292
/note="AluSg repeat: matches 2. .310 of consensus"
complement(27208. .27677)
/note="match: GSS: Em:AQ702871"
27819. .27957
/note="AluJb repeat: matches 163. .299 of consensus"
28701. .28930
/note="MLTIC repeat: matches 32. .301 of consensus"
29080. .29171
/note="MLTIIH repeat: matches 433. .526 of consensus"
29303. .29711
/note="L2 repeat: matches 2323. .2710 of consensus"
29712. .29834
/note="MLTIG repeat: matches 29. .147 of consensus"
30140. .30547
/note="L2 repeat: matches 1806. .2252 of consensus"
complement(30720. .31207)
/note="match: STS: Em:HS1J17T"
30919. .31120
/note="MER58A repeat: matches 1. .208 of consensus"
32109. .32472
/note="PHEIC repeat: matches 1. .371 of consensus"
33360. .33511
/note="MIR repeat: matches 98. .250 of consensus"
34031. .34331

```

```

/note="AluX repeat: matches 1. .297 of consensus"
34621. .34732
repeat_region
/note="MIR repeat: matches 120. .232 of consensus"
34888. .35007
repeat_region
/note="L2 repeat: matches 1836. .1959 of consensus"
35164. .35461
repeat_region
/note="AluSg repeat: matches 2. .298 of consensus"
35532. .35990
repeat_region
/note="L2 repeat: matches 2254. .2710 of consensus"
36631. .36790
repeat_region
/note="L2 repeat: matches 2342. .2501 of consensus"
37251. .37544
repeat_region
/note="AluX repeat: matches 1. .294 of consensus"
38446. .38493
repeat_region
/note="12 copies 4 mer caca 75% conserved"
38467. .38492
repeat_region

Query Match      99.9%; Score 1380.4; DB 9; Length 152966;
Best Local Similarity 99.9%; Pred. No. 6.8e-306;
Matches 1381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATTATGCTAACACATAAACATGCAGGGGTGGGAGCAGGGTCACAAAAGTGAGTGTG 60
Db 53422 CATTATGCTAACACATAAACATGCAGGGGTGGGAGCAGGGTCACAAAAGTGAGTGTG 53481
Qy 61 TCAATTCTACTTGAATGAAAGTTGAAATAATTTAAACAGTACGGGAAATGCAGAGCAA 120
Db 53482 TCAATTCTACTTGAATGAAAGTTGAAATAATTTAAACAGTACGGGAAATGCAGAGCAA 53541
Qy 121 TTTTCTCTCTGGTGACAATATAGTGTCCAACACTTGAAGTGATTTTAAAGATGTTTA 180
Db 53542 TTTTCTCTCTGGTGACAATATAGTGTCCAACACTTGAAGTGATTTTAAAGATGTTTA 53601
Qy 181 TTTAAATTAAAGAGTGATTTCCAAAGAAAATAAGGAAAAGAAAAGAAAAGT 240
Db 53602 TTTAAATTAAAGAGTGATTTCCAAAGAAAATAAGGAAAAGAAAAGAAAAGT 53661
Qy 241 AACGAAAACGAAAAGTATCAGTTTGTCTACTTAACCTTTGCAACGATACCTTTTATTT 300
Db 53662 AACGAAAACGAAAAGTATCAGTTTGTCTACTTAACCTTTGCAACGATACCTTTTATTT 53721
Qy 301 TCTTTAAGATTCCTGTTGTTTATACACAGATTTTAAAGTTTACTCTACTGTGCCCAAG 360
Db 53722 TCTTTAAGATTCCTGTTGTTTATACACAGATTTTAAAGTTTACTCTACTGTGCCCAAG 53781
Qy 361 TGAATTTCTTCTCCAGTCACAGTGTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTT 420
Db 53782 TGAATTTCTTCTCCAGTCACAGTGTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTT 53841
Qy 421 AGGGGCATCAATCACACCGAGAAGTCAAGCCCTCAACCTTCAACCTTCAACCTTCAACCTTCA 480
Db 53842 AGGGGCATCAATCACACCGAGAAGTCAAGCCCTCAACCTTCAACCTTCAACCTTCAACCTT 53901
Qy 481 GGATCTGATTTCTTCATATCAACCCACACATATAGGACACCTAAATGGTGGGGGTGG 540
Db 53902 GGATCTGATTTCTTCATATCAACCCACACATATAGGACACCTAAATGGTGGGGGTGG 53961
Qy 541 GGGACACGACTCACTTGAGTTTCTTGAAGGCTTCTGGCTTCCAGCCACGCAATTTGCC 600
Db 53962 GGGACACGACTCACTTGAGTTTCTTGAAGGCTTCTGGCTTCCAGCCACGCAATTTGCC 54021
Qy 601 CGCTCTGGATCTGCTTACGTTCCGGATTCGGTGCCAGTCCGGGGGTGTAGATGTTTC 660
Db 54022 CGCTCTGGATCTGCTTACGTTCCGGATTCGGTGCCAGTCCGGGGGTGTAGATGTTTC 54081
Qy 661 CTGACGGCCCCAAAGGTCCTGACGCGCGCGGTCACCTCTCTTACGGAAGACTTCGAAG 720
Db 54082 CTGACGGCCCCAAAGGTCCTGACGCGCGCGGTCACCTCTCTTACGGAAGACTTCGAAG 54141
Qy 721 CTGACACCTCTCTCTCATGATGACGACGCGCGCGCGCTAGAAAGGGTCCCCGTTTG 780
Db 54142 CTGACACCTCTCTCTCATGATGACGACGCGCGCGCGCTAGAAAGGGTCCCCGTTTG 54201

```



```
Db 792 GACGACGCGCGCCCGGTAGAGGGGTCCCGTTGCGGTACACAAAGCAGCTCTTCAC 733
Qy 804 GACGGCTGACAGAGTGGTGGACCTGGCGCTGCTCCGCTATCTTCGCCGCTGGCCG 863
Db 732 GACGGCTGACAGAGTGGTGGACCTGGCGCTGCTCCGCTATCTTCGCCGCTGGCCG 673
Qy 864 CCGCTCAGCTCGCTGCTCGCGTGGGAGGCACTCCGCTGCTCCCGGCGCTCACCGC 923
Db 672 CCGCTCAGCTCGCTGCTCGCGTGGGAGGCACTCCGCTGCTCCCGGCGCTCACCGC 613
Qy 924 ACCAGGCGCGGATCGCTCT 947
Db 612 ACCAGGCGCGGATCGCTCT 589

RESULT 6
AF181721/c 2167 bp mRNA linear PRI 10-JAN-2000
LOCUS Homo sapiens RU2S (RU2) mRNA, complete cds.
DEFINITION AF181721
ACCESSION AF181721
VERSION AF181721.1 GI:6684529
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2167)
AUTHORS Van Den Eynde, B.J., Gaugler, B., Probst-Keppler, M., Michaux, L., Devuyt, O., Lorgey, F., Weynants, P., and Boon, T.
TITLE A new antigen recognized by cytolytic T lymphocytes on a human kidney tumor results from reverse strand transcription
JOURNAL J. Exp. Med. 190 (12), 1793-1800 (1999)
MEDLINE 20069887
PUBMED 10601354
REFERENCE 2 (bases 1 to 2167)
AUTHORS Van den Eynde, B.J., Gaugler, B. and Pilotte, L.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-1999) Ludwig Institute for Cancer Research, Avenue Hippocrate, 74, Brussels 1200, Belgium
FEATURES
Source 1..2167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22.1"
1..2167
/gene="RU2"
303..1733
/gene="RU2"
/codon_start=1
/product="RU2S"
/protein_id="AAF23612.1"
/db_xref="GI:6684530"
/translation="MSGSSARSHLSQPVVKSVLVYRNGDPFFVAGRRVVIHERKVSF
EVFLKEVTGVQAFGAVRNITPRTGHRIRKLDIQSGNRYVAGGQEAFLKLYLDI
GEIKRPMVEVNTVEKVIHRSINVSARFRKPLQEPCTIFLIANGDLINPARRLLIPR
KTLNWDHVLQWYTEKTLRSGAVHRLYTLGLKIVSGEALENGQFVYVGRDKFKKL
PYGELLFDKSTMRRPFGKASLPPPIVGRSKSGNDRHSKSTVGSNDNSPQPLKR
KGRKEDVNSKLTKLQKNVLSQETIPNSDGEIFRAGAESETRGAEEVDEEDTQ
VEVPDVRPAEIVDEEEDGEKANKAEQDELFSGMNGDLEEGGREATDAPEQVEEIL
DHSQQAARPAVNGTDPEENGELQOYVNNELQLVLDKERKSQAGSGQDEADYDPPRP
PRPVKITSPSENNQNRDAAVA"
BASE COUNT 675 a 440 c 576 g 476 t
ORIGIN

Query Match 43.1%; Score 595.4; DB 9; Length 2167;
Best Local Similarity 99.8%; Pred. No. 5.7e-126;
Matches 596; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 554 ACTTGAGTTTCTGAGGCTTCCTGGCTCCAGCCACGTAATTCGCCCGCTCTGGATCT 613
Db 597 AATTGAGTTTCTTGAAGGCTTCCTGGCTCCAGCCACGTAATTCGCCCGCTCTGGATCT 538
```

```
Qy 614 GGCTAGCTTCCGGATTTCGGTGGCCAGTCCGCGGGTGTAGATGTTCTCTCAGCGCCCAA 673
Db 537 GGCTAGCTTCCGGATTTCGGTGGCCAGTCCGCGGGTGTAGATGTTCTCTCAGCGCCCAA 478
Qy 674 AGGTGCTCAAGCGCCGCGTCACTCCCTTCAGGAAGACTTCGAAGCTGGACACCTTCT 733
Db 477 AGGTGCTCAAGCGCCGCGTCACTCCCTTCAGGAAGACTTCGAAGCTGGACACCTTCT 418
Qy 734 TCTCATGGATGAGACGCGCGGCCCGCGTAGAAGGGTCCCGTTCCGGTTCACAAAGCA 793
Db 417 TCTCATGGATGAGACGCGCGGCCCGCGTAGAAGGGTCCCGTTCCGGTTCACAAAGCA 358
Qy 794 CGCTTTTCACAGCGGCTGAGACAGGTGGTGGACCTGGCGCTGCTCCGCTCATCTTCC 853
Db 357 CGCTTTTCACAGCGGCTGAGACAGGTGGTGGACCTGGCGCTGCTCCGCTCATCTTCC 298
Qy 854 CGCTGGCGCGCGCTCAGCTCGCTGCTCGGTCGCGGAGGCACTCCGCTGTCCCGAGCG 913
Db 297 CGCTGGCGCGCGCTCAGCTCGCTGCTCGGTCGCGGAGGCACTCCGCTGTCCCGAGCG 238
Qy 914 GCCTCACCGCACCCAGCGCGGGATCGCTTCTGAACGAACGAGAACTGACGAATCC 973
Db 237 GCCTCACCGCACCCAGCGCGGGATCGCTTCTGAACGAACGAGAACTGACGAATCC 178
Qy 974 ACAGGTGAAAGAGAAAGTAACGCGCGCTAGGCTGCACCCAGAGGAGACACTAGGA 1033
Db 177 ACAGGTGAAAGAGAAAGTAACGCGCGCTAGGCTGCACCCAGAGGAGACACTAGGA 118
Qy 1034 GCTTGAGGAGTCCGAGTACAGCTCAAGTTTTCACCGTGGCGTGACACCCCAATCAGG 1093
Db 117 GCTTGAGGAGTCCGAGTACAGCTCAAGTTTTCACCGTGGCGTGACACCCCAATCAGG 58
Qy 1094 ACCCGCAGTGGCGCACACACAGGTTCACCTGCTAGGGCAGANTCAAGGTGGAC 1150
Db 57 ACCCGCAGTGGCGCACACAGGTTCACCTGCTAGGGCAGANTCAAGGTGGAC 1

RESULT 7
AK027036/c 2003 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ23383 fis, clone HEP16466, highly similar to
DEFINITION AF181721 Homo sapiens RU2S mRNA.
ACCESSION AK027036
VERSION AK027036.1 GI:10440050
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
clone:HEP16466.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Ohtayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEO human cDNA sequencing project
Unpublished
TITLE 2 (bases 1 to 2003)
REFERENCE 2 (bases 1 to 2003)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Ohtayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
LOCATION/Qualifiers
```



```

source
1. .2003
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP18466"
/cell_line="HepG2"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"
misc_feature
1. .2003
/notes="highly similar to AF181721 Homo sapiens RU2S mRNA"
BASE COUNT 562 a 388 c 525 g 428 t
ORIGIN

Query Match 29.7%; Score 409.8; DB 9; Length 2003;
Best Local Similarity 99.5%; Pred. No. 2.4e-83;
Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 554 ACTTGAGTTTCTTGAAGGCTTCCTGGCTCCAGCAGCTAATTCGCCCGCTCTGGATCT 613
DB 413 AATTGAGTTTCTTGAAGGCTTCCTGGCTCCAGCAGCTAATTCGCCCGCTCTGGATCT 354

QY 614 GGTCTAGTTCGGATTCCGTGGCCAGTCCGCGGGGTAGATGTTCTTGACGCCCCAA 673
DB 353 GGTCTAGTTCGGATTCCGTGGCCAGTCCGCGGGGTAGATGTTCTTGACGCCCCAA 294

QY 674 AGGTGCTTGAACGCCGCGCTCACCCTTCAGAGAGCTTCGAAGCTGGACACCTTCT 733
DB 293 AGGTGCTTGAACGCCGCGCTCACCCTTCAGAGAGCTTCGAAGCTGGACACCTTCT 234

QY 734 TCTCATGATGACGACGCGCGCGCGCGGTAGAGAGGGTCCCGCTGGCGGTACACAAGCA 793
DB 233 TCTCATGATGACGACGCGCGCGCGCGGTAGAGAGGGTCCCGCTGGCGGTACACAAGCA 174

QY 794 CGCTCTTACGACGCGGTGACAGAGTGGCTGGACCTGGCGCTGCTGCCGTCTATCTTC 853
DB 173 CGCTCTTACGACGCGGTGACAGAGTGGCTGGACCTGGCGCTGCTGCCGTCTATCTTC 114

QY 854 CCGTGGCGCGCGCTCAGCTCGCTGCTTCGCGTGGGAGGACCTCCGCTGTCACGCG 913
DB 113 CCGTGGCGCGCGCTCAGCTCGCTGCTTCGCGTGGGAGGACCTCCGCTGTCACGCG 54

QY 914 GCCTCAGCGACCCAGGCGCGGGATCGCTCTCTGAACGAGCAACTGA 966
DB 53 GCCTCAGCGACCCAGGCGCGGGATCGCTCTCTGAACGAGCAACTGA 1

RESULT 8
AB032980/c 5548 bp mRNA linear PRI 10-MAY-2002
LOCUS Homo sapiens mRNA for KIAA1154 protein, partial cds.
DEFINITION AB032980
ACCESSION AB032980
VERSION AB032980.2 GI:20521781
KEYWORDS Homo sapiens adult male brain cDNA to mRNA, clone_lib:pBluescriptII
SOURCE SK plus clone:hh03679s1.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hirosewa.M., Nagase,T., Ishikawa,K., Kikuno,R., Nomura,N. and Ohara,O.
TITLE Characterization of cDNA clones selected by the GeneMark analysis from size-fractionated cDNA libraries from human brain
JOURNAL DNA Res. 6 (5), 329-336 (1999)
MEDLINE 20039618
PUBMED 10574461
REFERENCE 2 (bases 1 to 6548)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,

```

```

COMMENT
On May 9, 2002 this sequence version replaced yi:6330108.
FEATURES
source
1. .6548
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hh03679s1"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
/notes="This sequence was obtained by subcloning of the DNA fragments derived from two cDNA clones (1 - 1251 was derived from a RT-PCR product (Kidney) and 1252 - 6548 was derived from hh03679)."
1. .6548

```

gene

CDS

```

/gene="KIAA1154"
<1..1431
/notes="Start codon is not identified."
/sex="male"
/codon_start=1
/product="KIAA1154 protein"
/protein_id="BAA86468.2"
/db_xref="GI:20521782"
/translation="MSGSSARSHLSQPVVKSVLVYRNGDPFYAGRRVVIHEKVSSE
EVLKVTGQVQAPFGAVRNIVYPTGRHTRKLDIOISGNGYVAGGQEAFFKLNLYLDI
GEIKRPMVNVTEVPYIHSRINVSARFRKPLQEPCTIELIANGDLINPÁSRLLIPR
KTUNQWDHVLQWVTEKITLRSGAVHRDITLLEGLVSGALENGQFVAVGRDKFRKL
PYGELLFDKSTMRPFQKASSLPIVGSKSGNDRHSKTSVSGSSDSSPOPLKR
KGKEDVNSEKLTAKLVKLNQSETIPNSDEGIFKAGAESITRGAASQVEDQEDTQ
VEYVDQPAEIVDEEDGEKANKDAEQEDFSGMNGDLEEGREATDAPEQVEEIL
DHSEQOARPARVNGTDEENGELQOVNLEQLVLDKERKSGAGSGQDEADVDPQRP
PREVKITSPENENNOOKDYAQA"
BASE COUNT 2080 a 1124 c 1377 g 1967 t
ORIGIN

```

```

Query Match 21.2%; Score 293.4; DB 9; Length 6548;
Best Local Similarity 99.7%; Pred. No. 1.7e-56;
Matches 294; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 554 ACTTGAGTTTCTTGAAGGCTTCCTGGCTCCAGCAGCTAATTCGCCCGCTCTGGATCT 613
DB 295 AATTGAGTTTCTTGAAGGCTTCCTGGCTCCAGCAGCTAATTCGCCCGCTCTGGATCT 236

```

```

QY 614 GGTCTAGTTCGGATTCCGTGGCCAGTCCCGGGGTAGATGTTCTTGACGGCCCAA 673
DB 235 GGTCTAGTTCGGATTCCGTGGCCAGTCCCGGGGTAGATGTTCTTGACGGCCCAA 176

```

```

QY 674 AGGTGCTTGAACGCCGCGCTCACCCTCTCTGAGAGAGCTTCGAAGCTGGACACCTTCT 733
DB 175 AGGTGCTTGAACGCCGCGCTCACCCTCTCTGAGAGAGCTTCGAAGCTGGACACCTTCT 116

```

```

QY 734 TCTCATGATGACGACGCGCGCGCGGTAGAGAGGGTCCCGCTTCCGGTACACAAGCA 793
DB 115 TCTCATGATGACGACGCGCGCGCGGTAGAGAGGGTCCCGCTTCCGGTACACAAGCA 56

```

```

QY 794 CGCTCTTCAACGACGCGGTGAGACAGGTGGCTGGACCTGCGCTGCTGCGCTCAT 848
DB 55 CGCTCTTCAACGACGCGGTGAGACAGGTGGCTGGACCTGCGCTGCTGCGCTCAT 1

```

RESULT 9

AL589735

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

```

AL589735 19957 bp DNA linear ROD 29-JUN-2002
Mouse DNA sequence from clone RP23-211M15 on chromosome 13,
complete sequence.
AL589735
AL589735.12 GI:15795473
HTG.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```


repeat_region	5249. .5273	/rpt_family="GC-rich"	
repeat_region	5593. .6486	/rpt_family="SVA"	
repeat_region	6443. .6529	/rpt_family="GA-rich"	
repeat_region	6531. .6732	/rpt_family="L1"	
repeat_region	6733. .7033	/rpt_family="Alu"	
repeat_region	7034. .7249	/rpt_family="L1"	
repeat_region	7250. .7534	/rpt_family="Alu"	
repeat_region	7555. .8490	/rpt_family="L1"	
repeat_region	8528. .8726	/rpt_family="Achoho"	
misc_feature	8666. .8669	/note="match to EST BF314952 (NID:g11263123)"	
misc_feature	9031. .9360	/note="match to EST BF314952 (NID:g11263123)"	
misc_feature	9047. .9360	/note="match to EST BE259985 (NID:g9130920)"	
misc_feature	9088. .9089	/note="match to EST BF316717 (NID:g11265074)"	
misc_feature	9088. .9069	/note="match to EST BI198782 (NID:g14653803)"	
misc_feature	9144. .9147	/note="similar to Homo sapiens EST BF311850 (NID:g11259616)"	
misc_feature	9239. .9360	/note="match to EST BI198831 (NID:g14653852)"	
misc_feature	9239. .9360	/note="similar to Homo sapiens EST BI198308 (NID:g14653329)"	
misc_feature	9239. .9360	/note="similar to Homo sapiens EST BI198502 (NID:g14653523)"	
misc_feature	9240. .9360	/note="match to EST BF206235 (NID:g11099821)"	
misc_feature	9240. .9360	/note="match to EST BF311988 (NID:g11259763)"	
misc_feature	9240. .9360	/note="match to EST BI199896 (NID:g14654917)"	
misc_feature	9240. .9360	/note="similar to Homo sapiens EST BF313026 (NID:g11260952)"	
misc_feature	9313. .9360	/note="match to EST BF311981 (NID:g11259754)"	
misc_feature	9313. .9360	/note="similar to Homo sapiens EST BF316085 (NID:g11264380)"	
misc_feature	9313. .9360	/note="similar to Homo sapiens EST BI198293 (NID:g14653314)"	
misc_feature	9322. .9360	/note="match to EST BF316717 (NID:g11265074)"	
misc_feature	9322. .9360	/note="match to EST BI198782 (NID:g14653803)"	
misc_feature	9322. .9360	/note="similar to Homo sapiens EST BF311850 (NID:g11259616)"	
misc_feature	9333. .9360	/note="match to EST BE313758 (NID:g9134050)"	
misc_feature	9338. .9360		
Query Match	6.28;	Score 85.4;	DB 9; Length 139357;
Best Local Similarity	59.18;	Pred. No. 2e-08;	
Matches 165; Conservative	0;	Mismatches 111;	Indels 3; Gaps 1;
Qy	526	ATGGGTGGCGGTGGGGAGACCGACTGACCTTGAGTTCTTGAAGGCTTCCTGGCCTCCA	585

Db 118260 AGGCGCGTGGGGGCGGGCGGCACCTTACTCGAGCTCCTTGAAGCGTCCGGCGCCGCC 118201

Qy 586 GCCACGTAAATGGCCCCCTCTGGATCTGGTCTAGCTTCCGGATTCGGTGGCCAGTCCGC 645

Db 118200 GCCAGTACTTGGCGCGCTGAGCGGTCCAGCCCGACCGCGTGGCCCGCGTG 118141

Qy 646 GGGGTGTAGATGTTCTTCTGACGGCCCAAGGCTGCTGAACGCGCCGCCCTCACCTCCCTTC 705

Db 118140 GCGGTGAAGAGCGCGCGCACG---CCGAACGGGAGCTCCACCTGCTCCGTGAGTGCCTCC 118084

Qy 706 AGGAAGACTTGAAGCTGGACACCTTCTCTCATGATGACGACGCGCGCCCGGTAG 765

Db 118083 AGCAGCGCTCGAAGTGGCGCGCGCGCGACACGACGAACTTCTTCCCGACGTAG 118024

Qy 766 AAGGGTCCCGTTCGGGTACACAGACGCTCTTCACG 804

Db 118023 AACGGTCCCGTTCGGGTACACGATGCTCTTGGCG 117985

RESULT 12

HSDJ622L5/c

LOCUS HSDJ622L5 115756 bp DNA linear PRI 22-NOV-2001

DEFINITION Human DNA sequence from clone RP4-622L5 on chromosome 1p34.2-36.11, complete sequence.

ACCESSION AL049795

VERSION AL049795.21 GI:17065925

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 115756)

AUTHORS Frankland,J.

TITLE Direct Submission

JOURNAL Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 25, 2001 this sequence version replaced gi:6010175. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chrl This sequence is the entire insert of clone RP4-622L5 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-622L5 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTORS: pCYPAC2.

FEATURES

source

Location/Qualifiers

1. .115756

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="p34.2-36.11"

/clone="RP4-622L5"

/clone_lib="RPCI-4"

```
repeat_region 1645. .1810
/notes="TIGER1 repeat: matches 2259. .2417 of consensus"
repeat_region 2123. .2437
/notes="TIGER1 repeat: matches 1957. .2259 of consensus"
repeat_region 2726. .2782
/notes="TIGER1 repeat: matches 1904. .1957 of consensus"
repeat_region 3067. .3486
/notes="TIGER1 repeat: matches 1501. .1904 of consensus"
repeat_region 3781. .4180
/notes="TIGER1 repeat: matches 1117. .1501 of consensus"
repeat_region 5033. .5082
/notes="TIGER1 repeat: matches 1090. .1117 of consensus"
repeat_region 5372. .5501
/notes="TIGER1 repeat: matches 966. .1090 of consensus"
repeat_region 5628. .5704
/notes="TIGER1 repeat: matches 893. .966 of consensus"
repeat_region 5991. .6492
/notes="TIGER1 repeat: matches 411. .893 of consensus"
repeat_region 6800. .6835
/notes="TIGER1 repeat: matches 378. .411 of consensus"
repeat_region 6968. .7022
/notes="TIGER1 repeat: matches 1. .378 of consensus"
repeat_region 8281. .8358
/notes="3 copies 26 mer 79% conserved"
repeat_region 8301. .8360
/notes="30 copies 2 mer ta 81% conserved"
repeat_region 17674. .17777
/notes="4 copies 26 mer 77% conserved"
repeat_region 17685. .17784
/notes="50 copies 2 mer at 82% conserved"
misc_feature 19385. .19542
/notes="match: GSS: Em:AQ306911"
repeat_region 19551. .19574
/notes="12 copies 2 mer tt 100% conserved"
repeat_region 20865. .20922
/notes="29 copies 2 mer at 70% conserved"
repeat_region 20934. .21043
/notes="55 copies 2 mer at 70% conserved"
repeat_region 20940. .21043
/notes="4 copies 26 mer 74% conserved"
repeat_region 21168. .21234
/notes="tRNA-Asp-GAC repeat: matches 1. .75 of consensus"
misc_feature 28263. .28691
/notes="match: GSS: Em:AQ317905"
misc_feature complement(29964. .30447)
/notes="match: GSS: Em:AQ732256"
misc_feature 30465. .30718
/notes="match: STS: Em:G58909
match: GSS: Em:AQ376964"
repeat_region 38923. .39446
/notes="L1MPC repeat: matches 1546. .2095 of consensus"
repeat_region 39761. .39882
/notes="L1MPC repeat: matches 1429. .1546 of consensus"
misc_feature complement(43254. .43726)
/notes="match: GSS: Em:AQ146371"
repeat_region 45555. .46050
/notes="MER44C repeat: matches 1. .728 of consensus"
misc_feature 46055. .46468
/notes="match: GSS: Em:AQ404362"
repeat_region 48914. .48982
/notes="MER21B repeat: matches 718. .790 of consensus"
repeat_region 48985. .49173
/notes="MER21B repeat: matches 77. .263 of consensus"
misc_feature 49285. .49651
/notes="match: GSS: Em:AQ761532"
misc_feature 52462. .53177
/notes="match: GSS: Em:AQ404082"
misc_feature 54716. .54865
/notes="match: GSS: Em:AQ680771"
repeat_region 60551. .60614
/notes="32 copies 2 mer tt 70% conserved"
misc_feature complement(62219. .62382)
/notes="match: STS: Em:G20585"
```

```
misc_feature complement(66030. .66267)
/notes="match: GSS: Em:AQ514287"
misc_feature complement(66041. .66267)
/notes="match: GSS: Em:AQ803288"
misc_feature 66066. .66255
/notes="match: GSS: Em:B69239"
misc_feature 68882. .69541
/notes="CpG island"
/notes="evidence=not_experimental"
repeat_region 71427. .71448
/notes="11 copies 2 mer tt 100% conserved"
misc_feature 73527. .74249
/notes="match: GSS: Em:AQ892680"
misc_feature 74314. .74926
/notes="match: GSS: Em:AQ441232"
misc_feature 74532. .74981
/notes="match: GSS: Em:AQ812740"
misc_feature 74626. .75008
/notes="match: GSS: Em:AQ091614"
misc_feature complement(74682. .75020)
/notes="match: STS: Em:L29744"
misc_feature complement(74683. .74951)
/notes="match: STS: Em:L29729 Em:L29763"
misc_feature 74717. .74867
/notes="match: STS: Em:G15744"
misc_feature 82892. .83089
/notes="match: STS: Em:G09447"
misc_feature 83391. .83481
/notes="match: GSS: Em:A2310334"
repeat_region 85284. .85337
/notes="27 copies 2 mer tt 88% conserved"
misc_feature 85339. .85718
/notes="match: STS: Em:G28385"
misc_feature complement(87250. .87612)
/notes="match: STS: Em:G06968"
misc_feature complement(87455. .87579)
/notes="match: STS: Em:G26511"
misc_feature complement(87851. .88313)
/notes="match: GSS: Em:AQ536603"
repeat_region 89634. .89685
/notes="16 copies 2 mer tc 87% conserved"
misc_feature 91856. .92386
/notes="match: GSS: Em:B85187"
repeat_region 92101. .92152
/notes="26 copies 2 mer tt 71% conserved"
misc_feature 94297. .95612
/notes="CpG island"
/notes="evidence=not_experimental"
misc_feature complement(98493. .98894)
/notes="match: GSS: Em:AQ207463"
misc_feature 103719. .104155
/notes="match: GSS: Em:AQ543924"
repeat_region 104406. .104437
/notes="MER57B repeat: matches 303. .334 of consensus"
repeat_region 107785. .107841
/notes="L1MB1 repeat: matches 6111. .6171 of consensus"
misc_feature 111249. .111767
/notes="CpG island"
/notes="evidence=not_experimental"
misc_feature 113362. .113450
/notes="match: STS: Em:G16146
match: GSS: Em:AQ806873"
BASE COUNT 29576 a 26665 c 27977 g 31538 t
ORIGIN
```

Query Match 5.5%; Score 75.8; DB 9; Length 115756;
Best Local Similarity 58.2%; Pred.No.3e-06;
Matches 153; Conservative 0; Mismatches 107; Indels 3; Gaps 1;

QY 547 CCGACTCACTTGAGTTTCTTGAAGGCTTCTGGCTCCAGCCACGTAATTCGCCCGCTC 606
DB 98696 CCACACTGAGCTTGGAATCGTTCAATCCAGCGGCACACTGCTCTGTTTC 98637

QY 607 TGGATCTGCTAGTTCCTCCGATTCGGTGGCCAGTCCCGGGGTGTAGATGTTCTGTACG 666
Db 98636 TTCAAGTCTGCCAGGTGCTGACAGGTGGCCATGACAAAGGTGTGTAGAGGCGCAC 98577
QY 667 GCCCAAGAGGTGCTGAACCGCCGCTCACTTTCAGGAAGACTTCAAGCTGCAC 726
Db 98576 GCC---AGTGGGGCTGCACAGCTGATGTCACTTCGAGAGGAAGGCTCCATGTTGGGG 98520
QY 727 ACCTTCTTCATGATGACGACGCGCGCCCGCTGAGAAGGGTCCCGCTTGGGTAC 786
Db 98519 AAGCGGGCTTGAGTCAACACAGCTGGGAGCGCTGGGAAGATGGTCCCATTCGGTAC 98460
QY 787 ACAGCAGCTTTCACACGCG 809
Db 98459 ACCACTACCTCTTGGCTGCTGG 98437

RESULT 13
HSDJ675E8/c
LOCUS HSDJ675E8 322840 bp DNA linear HTG 17-JUL-2002
DEFINITION Homo sapiens chromosome 1 clone RP4-675E8 map p34.1-35.3, ***
SEQUENCING IN PROGRESS ***, 45 unordered pieces.
ACCESSION AL121991
VERSION AL121991.41 GI:21912433
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 322840)
Matthews.N.
Direct Submission
Submitted (15-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Jul 19, 2002 this sequence version replaced gi:21535828.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: dj675E8
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator ABI; 0% of reads
Chemistry: Dye-terminator; 8% of reads
Chemistry: Dye-terminator ET-amersham; 4% of reads
Dye-terminator Big Dye; 86% of reads
Consensus quality: 309409 bases at least Q40
Consensus quality: 313127 bases at least Q30
Consensus quality: 315404 bases at least Q20
Insert size: 318440; sum-of-contigs
Insert size: 128138; 23.4% error; agarose-fp
Quality coverage: 6.9% in Q20 bases; sum-of-contigs Quality
coverage: 22.25x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 45 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 54147: contig of 54147 bp in length
* 54148 54247: gap of 100 bp
* 54248 71025: contig of 16778 bp in length
* 71026 71125: gap of 100 bp
* 71126 75812: contig of 4687 bp in length
* 75813 75912: gap of 100 bp
* 75913 116848: contig of 40936 bp in length

* 116849 116948: gap of 100 bp
* 116949 118963: contig of 2015 bp in length
* 118964 119063: gap of 100 bp
* 119064 125711: contig of 6548 bp in length
* 125612 125711: gap of 100 bp
* 125712 164112: contig of 38401 bp in length
* 164113 164212: gap of 100 bp
* 164213 180038: contig of 15826 bp in length
* 180039 180138: gap of 100 bp
* 180139 184418: contig of 4280 bp in length
* 184419 184518: gap of 100 bp
* 184519 187132: contig of 2614 bp in length
* 187133 187232: gap of 100 bp
* 187233 190323: contig of 3091 bp in length
* 190324 190423: gap of 100 bp
* 190424 192746: contig of 2323 bp in length
* 192747 192846: gap of 100 bp
* 192847 194893: contig of 2047 bp in length
* 194894 194993: gap of 100 bp
* 194994 197169: contig of 2176 bp in length
* 197170 197269: gap of 100 bp
* 197270 199729: contig of 2460 bp in length
* 199730 199829: gap of 100 bp
* 199830 202634: contig of 2805 bp in length
* 202635 202734: gap of 100 bp
* 202735 205712: contig of 2978 bp in length
* 205713 205812: gap of 100 bp
* 205813 208288: contig of 2476 bp in length
* 208289 208388: gap of 100 bp
* 208389 213338: contig of 4950 bp in length
* 213339 213438: gap of 100 bp
* 213439 216918: contig of 3480 bp in length
* 216919 217018: gap of 100 bp
* 217019 220931: contig of 3913 bp in length
* 220932 221031: gap of 100 bp
* 221032 224098: contig of 3067 bp in length
* 224099 224198: gap of 100 bp
* 224199 226367: contig of 2169 bp in length
* 226368 226467: gap of 100 bp
* 226468 229613: contig of 3146 bp in length
* 229614 229713: gap of 100 bp
* 229714 232104: contig of 2391 bp in length
* 232105 232204: gap of 100 bp
* 232205 265235: contig of 33031 bp in length
* 265236 265335: gap of 100 bp
* 265336 268023: contig of 2688 bp in length
* 268024 268123: gap of 100 bp
* 268124 270219: contig of 2096 bp in length
* 270220 270319: gap of 100 bp
* 270320 278019: contig of 7700 bp in length
* 278020 278119: gap of 100 bp
* 278120 280129: contig of 2010 bp in length
* 280130 280229: gap of 100 bp
* 280230 282545: contig of 2316 bp in length
* 282546 282645: gap of 100 bp
* 282646 285394: contig of 2749 bp in length
* 285395 285494: gap of 100 bp
* 285495 287702: contig of 2208 bp in length
* 287703 287802: gap of 100 bp
* 287803 289871: contig of 2069 bp in length
* 289872 289971: gap of 100 bp
* 289972 293367: contig of 3396 bp in length
* 293368 293467: gap of 100 bp
* 293468 295482: contig of 2015 bp in length
* 295483 295582: gap of 100 bp
* 295583 299991: contig of 4409 bp in length
* 299992 300091: gap of 100 bp
* 300092 302207: contig of 2116 bp in length
* 302208 302307: gap of 100 bp
* 302308 304990: contig of 2683 bp in length
* 304991 305090: gap of 100 bp
* 305091 307235: contig of 2245 bp in length
* 307236 307435: gap of 100 bp

* 307436 309641: contig of 2206 bp in length
 * 309642 309741: gap of 100 bp
 * 309742 315531: contig of 5790 bp in length
 * 315532 315631: gap of 100 bp
 * 315632 318321: contig of 2690 bp in length
 * 318322 318421: gap of 100 bp
 * 318422 320469: contig of 2048 bp in length
 * 320470 320569: gap of 100 bp
 * 320570 322840: contig of 2271 bp in length.

FEATURES

```

location:"Quadrilateris
1. .322840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p34.1-35.3"
/clone="RP4-675P8"
/clone_lib="RPC1-4"
1. .34147
/note="assembly_fragment:11504
fragment_chain:1"
54248. .71025
/note="assembly_fragment:09814
fragment_chain:1"
71126. .75812
/note="assembly_fragment:11398
fragment_chain:1"
75913. .116848
/note="assembly_fragment:03358
fragment_chain:1"
116949. .118963
/note="assembly_fragment:10591
fragment_chain:2"
119064. .125611
/note="assembly_fragment:09366
fragment_chain:2"
125712. .164112
/note="assembly_fragment:11336
fragment_chain:3"
164213. .180038
/note="assembly_fragment:02790
fragment_chain:3"
180139. .184418
/note="assembly_fragment:00167"
184519. .187132
/note="assembly_fragment:00622"
187233. .190323
/note="assembly_fragment:01142"
190424. .192746
/note="assembly_fragment:01318"
192847. .194893
/note="assembly_fragment:01614"
194994. .197169
/note="assembly_fragment:01627"
197270. .199729
/note="assembly_fragment:01811"
199830. .202634
/note="assembly_fragment:01839"
202735. .205712
/note="assembly_fragment:02187"
205813. .208288
/note="assembly_fragment:02212"
208389. .213338
/note="assembly_fragment:02545"
213439. .216918
/note="assembly_fragment:02573"
217019. .220931
/note="assembly_fragment:02907"
221032. .224098
/note="assembly_fragment:04499"
224199. .226367
/note="assembly_fragment:05011"
226468. .229613
/note="assembly_fragment:05130"

```


RESULT 15

AB047038/c
LOCUS AB047038 3295 bp mRNA linear INV 11-JAN-2002
DEFINITION Halocynthia roretzi Hrdoublecortin mRNA, complete cds.
ACCESSION AB047038
VERSION AB047038.1 GI:13537215
KEYWORDS
SOURCE Halocynthia roretzi cDNA to mRNA.
ORGANISM Halocynthia roretzi
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
Stolidobranchia; Pyuridae; Halocynthia.
1
Yagi, K. and Makabe, K.W.
Neural marker genes differently expressed in subsets of embryonic
neural cells of the ascidian Halocynthia roretzi
Unpublished
2 (bases 1 to 3295)
Yagi, K.
Direct Submission
Submitted (08-AUG-2000) Kasumi Yagi, Kyoto University, Department
of Zoology, Graduate School of Science, Kitashirakawa Oiwake-cho,
Sakyo-ku, Kyoto-shi, Kyoto 606-8502, Japan
(E-mail: kasumi@ascidian.zool.kyoto-u.ac.jp,
URL: http://hoya.zool.kyoto-u.ac.jp/FRINDEX.HTM,
Tel: 81-075-753-4082 (ex. 4082), Fax: 81-075-705-1113)
Tel: 81-075-753-4082 (ex. 4082), Fax: 81-075-705-1113)
Location/Qualifiers
1..3295
/organism="Halocynthia roretzi"
/db_xref="taxon:7729"
1..3295
/gene="Hrdoublecortin"
129..2540
/gene="Hrdoublecortin"
/codon_start=1
/protein_id="BAB40784.1"
/db_xref="GI:13537215"
/translation="MTLVHAQQDHNHMHOPETAKNVHVVYNGDQYYPGRKFVYVNRRIY
CDFGLNLTSLGKPSFGAVRIYTPNAGHRVNLHEINNNHVI VAGGAERERKLHY
QNTPLPRMPKPKHSYSTRPVYHNSKMNVPARKOLVOGPCIYVYGNRDINGVAI
KMLLSARIKMWDMVLEETEKISIRTGQAVRLYTLDGVLIOGSENLENGRYVAVG
YERFKAPGGERAQT LAMLSRPHVTLPPIKPKALNSKPAEPNQSQAQYAKKK
QMEEMTRQPPNLTPLNGERKNRPHKPIPPPPSPSESDSVFHAKPVKVRKSGASGRVT
QVYPSVDDDDDDVYRAKSEYKAEVAEDEDTKVDLPVQMPADTIDDEDLDYVQHE
EAMNVKQERSPLNDQOTIMQADENAFKPSKPSPSHSITTESEGEVNTDEKRS
NVTTPSPSRINREDEKRENAALALOSTYRGLVRRDLKKKYSTVPIAEOKVET
QDLVIEKALNLEQPLAVGSENPSEDLVISENQENONPVILDSLSQOAEATPS
QESKNDDGDDDDDDGKAPKPNPVLGQDSPNLINSIDINDIAAKSDDLVAKEDD
DVTSEENAKVDVDSRHVTPLAAKSETHDKNPEDEESDKHNDVTEPVFDLUSGQAVON
LPSNDVNESNLNEQPNEDNLPDADGTESEGRDTRKSQSESTNLAHASSII
DNSNEEGEKORSGEESQEISNTPSARPATSESTENTPDQTPGKRIDEQPIPNQPTS
LTEQV"

FEATURES

Source

gene

CDS

BASE COUNT 1153 a 651 c 680 g 811 t
ORIGIN

Query Match 4.9%; Score 67.8; DB 3; Length 3295;
Best Local Similarity 52.3%; Pred. No. 9.8e-05;
Matches 150; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
Qy 517 GGCACCTAATGGTGGCGGTGGGGAGACCGACTACTTGAGTTCTTGAAGGCTTCC 576
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 469 GGCATCGGTTTCTTGGAGTGGAGTGTGTTTGATAATGAAGTTTCTGAATCTTCT 410
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 577 TGGCCTCCAGCCAGTAATTCGCCCGCTCTGGATCGGTCTAGCTTCCGGATTCGGTGG 636
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 409 GCTCCGGCGGCAACAATACCATGTTTATGTTAATTCGGAAGTTCGTTAACCCGATGT 350
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 637 CCAGTCCCGGGGTGTAGATGTTCTTACGCGCCCAAGAGGTGCTGAACGCCCGCGTC 696
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 349 CCCCGTTGGTGTGTAGATGTTCCGCACACGCCCGGAAGACGGTTTCAGGCCGCTTGA 290
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Qy 697 ACCTCTTCAGGAAGATTTCGAAGCTGGACACCTTCTTCTCATGGATGACGACGCGGC 756
||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:07:35 ; Search time 60 Seconds
(without alignments)
7063.788 Million cell updates/sec

Title: US-09-674-593-1
Perfect score: 1382
Sequence: 1 cattatgctaacagcataaa.....taagcaattaagtctctgg 1382

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
- score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	46.8	3.4	7218	1	US-08-232-463-14
C 2	44.8	3.2	1212	4	US-09-182-145-34
C 3	44.8	3.2	1212	4	US-09-182-145-35
C 4	43.4	3.1	8438	1	US-07-945-283-1
C 5	43	3.1	1519	1	US-07-971-759-19
C 6	42.2	3.1	1441	4	US-08-821-994-63
C 7	42.2	3.1	3238	4	US-08-123-934A-5
C 8	42.2	3.1	3238	5	PCT-US94-10080-5
C 9	40.2	2.9	2581	4	US-09-363-708-1
C 10	39.8	2.9	4403765	4	US-09-103-840A-2
C 11	39.4	2.9	688	6	5498694-3
C 12	39.4	2.9	2648	4	US-09-417-455-6
C 13	39.4	2.9	2648	4	US-09-348-942-6
C 14	39.4	2.9	2648	4	US-09-457-626-6
C 15	39.2	2.8	732	4	US-09-149-476-66
C 16	39	2.8	746	4	US-09-013-810-1
C 17	39	2.8	1176	4	US-09-372-422A-25
C 18	38.8	2.8	1931	3	US-09-019-942-2
C 19	38.8	2.8	1931	4	US-09-099-041A-1
C 20	38.8	2.8	1931	4	US-09-245-281-1
C 21	38.8	2.8	1931	4	US-09-470-271-2
C 22	38.8	2.8	1931	4	US-09-207-359B-1
C 23	38.8	2.8	80246	4	US-09-078-294-4
C 24	38.8	2.8	80595	4	US-09-078-294-3
C 25	38.6	2.8	1114	4	US-09-152-060-41
C 26	38.4	2.8	1740	4	US-09-125-642C-2
C 27	38.4	2.8	1742	4	US-09-125-642C-13

Sequence 1, Appl1	28	38.4	2.8	2239	4	US-09-196-390-1
Sequence 8, Appl1	C 29	38.4	2.8	5515	4	US-09-125-642C-8
Sequence 12, Appl1	C 30	38.4	2.8	5519	4	US-09-125-642C-12
Sequence 16, Appl1	C 31	38.4	2.8	9468	1	US-08-325-547-10
Sequence 2, Appl1	C 32	38.2	2.8	1738	2	US-08-379-482A-2
Sequence 11, Appl1	C 33	38	2.7	948	2	US-08-924-759-11
Sequence 4, Appl1	C 34	38	2.7	948	3	US-09-248-335-11
Sequence 4, Appl1	C 35	38	2.7	1046	1	US-08-361-467B-4
Sequence 4, Appl1	C 36	38	2.7	1046	1	US-08-484-332C-4
Sequence 3, Appl1	C 37	38	2.7	2760	1	US-08-101-593-3
Sequence 3, Appl1	C 38	37.6	2.7	1364	1	US-08-265-087-3
Sequence 3, Appl1	C 39	37.6	2.7	1364	1	US-08-621-493-3
Sequence 3, Appl1	C 40	37.6	2.7	1364	2	US-08-965-688-3
Sequence 3, Appl1	C 41	37.6	2.7	1364	4	US-09-260-173-3
Sequence 7, Appl1	C 42	37.6	2.7	1374	2	US-08-758-621-7
Sequence 7, Appl1	C 43	37.6	2.7	1374	4	US-09-107-858-7
Sequence 13, Appl1	C 44	37.6	2.7	2448	4	US-08-487-596-13
Sequence 9, Appl1	C 45	37.6	2.7	2450	2	US-08-466-589-9

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pz9ptc-fls
US-08-232-463-14

Query Match 3.4%; Score 46.8; DB 1; Length 7218;

```
Best Local Similarity 2.2%; Pred. No. 0.018;
Matches 6; Conservative 169; Mismatches 101; Indels 0; Gaps 0;

Qy 10 AACAGCATAAACATCGAGGGGTGGGAGCGGTCACAAAAGTGAAGTCTTGTCAATTCTA 69
Db 1334 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1275
Qy 70 CTGGAATGAAGGTTGAATTAATTAACAGTACGGGAATGCAGAGCAATTTCTCTCT 129
Db 1274 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1215
Qy 130 CTGGTGACAATAGTGTCCAACTGCGAAGTCAATTTTAAAGATGTTTAAATTA 189
Db 1214 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1155
Qy 190 AAAGATCGATTTCACGAGAAATAAGGAAAGGAAAGAAAGAAAGAAAGAAAGAA 249
Db 1154 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1095
Qy 250 CGCAAAGTATCATGTTGGTCACTAACCTTTGCAAG 285
Db 1094 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059

RESULT 2
US-09-182-145-34
; Sequence 34, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 34
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-34

Query Match 3.2%; Score 44.8; DB 4; Length 1212;
Best Local Similarity 55.0%; Pred. No. 0.024;
Matches 88; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 96 AAACAGTACGGGAATGCAGAGCAATTTCTCTCTGTGACAATATAGTGTCCAACT 155
Db 181 AAATGTCAGAGAACCTGGAGATATATTTCTGAGCTCAAGATTTCTGAAAACCAAGCAA 122
Qy 156 TGGAGTGTATTTTAAAGATGTTTAAATTAAGAGTGAATTCACAGGAAAAAA 215
Db 121 TGGGGGAAAAGTTAGTCAATCTGTCATATATAAAAAATAGTCAGTAAAAA 62
Qy 216 ATAAGGAAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 255
Db 61 AAAAAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 22

RESULT 4
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LLT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
```

```
RESULT 3
US-09-182-145-35/c
; Sequence 35, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 35
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-35

Query Match 3.2%; Score 44.8; DB 4; Length 1212;
Best Local Similarity 55.0%; Pred. No. 0.024;
Matches 88; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 96 AAACAGTACGGGAATGCAGAGCAATTTCTCTCTGTGACAATATAGTGTCCAACT 155
Db 181 AAATGTCAGAGAACCTGGAGATATATTTCTGAGCTCAAGATTTCTGAAAACCAAGCAA 122
Qy 156 TGGAGTGTATTTTAAAGATGTTTAAATTAAGAGTGAATTCACAGGAAAAAA 215
Db 121 TGGGGGAAAAGTTAGTCAATCTGTCATATATAAAAAATAGTCAGTAAAAA 62
Qy 216 ATAAGGAAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 255
Db 61 AAAAAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 22

RESULT 4
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LLT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
```

```
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P.
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8438 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 622..6495
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1099, "g")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1267, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1381, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1566, "c")
; NAME/KEY: variation
; LOCATION: replace(7010, "g")
; US-07-945-283-1

Query Match 3.1%; Score 43.4; DB 1; Length 8438;
Best Local Similarity 57.9%; Pred. No. 0.16; Mismatches 56; Indels 0; Gaps 0;
Matches 77; Conservative 0;

Qy 790 AGCAGCTCTTTCAGAGCGGCTGAGACAGGTGGCTGGACCTGGCGCTGCTCGCTCATC 849
Db 6349 AGCACTGCAGCGCAGAGCGCGCGCGCGCGCGCGCGCGCTCGTCCCTCGC 6290

Qy 850 TTCCCGCTGCGCGCGCTCAGCTGCTGCTGCTGCGTGGGAGGACCTCCGCTGTC 909
Db 6289 AGCCTCCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6230

Qy 910 AGCGGCTCACCG 922
Db 6229 GCGCGCGCGCGCG 6217

RESULT 5
US-07-971-759-19/c
; Sequence 19, Application US/07971759
; Patent No. 5573943
; GENERAL INFORMATION:
; APPLICANT: Saul, Alan J.
; APPLICANT: Cooper, Juan A.
; APPLICANT: Irving, David O.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF RHOPTRY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
```

```
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/971.759
; FILING DATE: 01-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU91/00338
; FILING DATE: 01-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK 1525 90
; FILING DATE: 02-AUG-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 61..1254
; US-07-971-759-19

Query Match 3.1%; Score 43; DB 1; Length 1519;
Best Local Similarity 54.9%; Pred. No. 0.082;
Matches 107; Conservative 0; Mismatches 85; Indels 3; Gaps 1;

Qy 150 AACACTTGGAGTGTATTTTAAAGATGTTTATTTTAAATTTAAAGATGATGATTTTCAAGGA 209
Db 1449 AATATTTAAATGTCAACTTAAAGAAATATTTAAACAAAAAATAATTTAAATTTAAATTTAA 1390

Qy 210 AAAAAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 266
Db 1389 AATAAATAACAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1330

Qy 267 GGTCACTAACCTTTTCAAGGATACCTTTTATTTTCTTTAAGATTCCTGTTGTTTATACA 326
Db 1329 TATCAGAAACCTTTTAAATATATAAATGCTATTTTGTAAATAGTTAATGTTTATTTATA 1270

Qy 327 CAGATTTTAAAGTTTA 341
Db 1369 ATGTATTTTATTTTA 1255

RESULT 6
US-08-821-994-63
; Sequence 63, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier RP
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821.994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1441
```


Qy 171 AGAATGTTTATTTAAATTAAGGATGATTTCCAGGAGAAAAAATAAGGAAAGGAAA 230
Db 3141 AATAACTGTTTCAAGTCAAAAAATAAGGATGATTTCCAGGAGAAAAAATAAGGAAAGGAAA 3200
Qy 231 GAAAAAACTGAACAGAAA 249
Db 3201 AAAAAAATAAAAAAAA 3219

RESULT 9
US-09-363-708-1
; Sequence 1, Application US/09363708
; Patent No. 6399747
; GENERAL INFORMATION:
; APPLICANT: Schmandt, et al.
; TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/363,708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 01017/34451
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; DESCRIPTION: /desc = "hPAL cDNA"

Query Match 2.9%; Score 40.2; DB 4; Length 2581;
Best Local Similarity 64.5%; Pred. No. 0.61;
Matches 60; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 164 ATTTTAAAGATGTTTATTTAAATTAAGGATGATTTCCAGGAGAAAAAATAAGGAA 223
Db 2479 ATTTCATAATACCTGCTTAAATAAAGGCTTTATGATTCGGGAAAAAATAAGGAA 2538

Qy 224 AAGGAAGAAAAAAGCTGACAGAAAAACGCAAAA 256
Db 2539 AAGGAAAAAATAAGGAAAAAATAAGGAAAAAATAAGGAAAAAATAAGGAAAA 2571

RESULT 10
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 2.9%; Score 39.8; DB 4; Length 4403765;
Best Local Similarity 47.1%; Pred. No. 31;
Matches 154; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

Qy 654 GATGTTCTGTACGCGCCCAAGGTTGCTGAACGCGCGGTCACTCTCTTCAGGAAGAC 713
Db 2059263 GACATTAGTCCGGTGCCTGCGCGCCGCGGACCGCCGCGGACCAACACCGCC 2059332

Qy 714 TTCGAAGCTGGACACCTTCTTCATGATGACGACGCGCGCGCGCGTGAAGAGGGTTC 773
Db 2059323 GTCTCGCGCGCGCGCGGTCACTGCGCGGTGCGCGGAGTCCGCGGACGCGCGCGCC 2059382

Qy 774 CCCGTTGGGTTACACAGCAGCTCTTCACGACGCGGTGACACAGCTGGCTGGACCTGCG 833
Db 2059383 GCGCGCGCGGACAGCAGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2059442

Qy 834 GCTG-CTCGCGCTCATCTTCCCGCTGCGCGCGCGCTCAGCTCGCTGCTCGCTCGGGA 892
Db 2059443 GTGCGCGGACGCTAGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2059502

Qy 893 GGCACCTCCGCTGTCCAGCGCGCTCACCGCACCGCGCGCGCGCGCGCGCGCGCGCGCG 952
Db 2059503 GGCTCGCGCGTTCGCGCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2059562

Qy 953 GAACGAGAACTGAGGAATCCACAGGT 979
Db 2059563 GATCAAGATTCCGCGCGCGCGCGGT 2059589

RESULT 11
5498694-3
; Patent No. 5498694
; APPLICANT: RUOSLAHTI, ERKKI I.
; TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
; INTEGRIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,967
; FILING DATE: 10-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 973,547
; FILING DATE: 09-NOV-1992
; APPLICATION NUMBER: 357,824
; FILING DATE: 25-MAY-1989
; SEQ ID NO:3
; LENGTH: 688
5498694-3

Query Match 2.9%; Score 39.4; DB 6; Length 688;
Best Local Similarity 59.3%; Pred. No. 0.49;
Matches 67; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 144 GTGTCCAACTTGAAGTGAATTTTAAAGATGTTTAAATTAAGGATGATTC 203
Db 564 GTCTCAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAAGTTACTTCCA 623
Qy 204 CAAGAAAAAATAAGGAAAAAATAAGGAAAAAATAAGGAAAAAATAAGGAAAAAATAAGG 256


```
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 2648
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-626-6

Query Match      2.9%   Score 39.4;   DB 4;   Length 2648;
Best Local Similarity 52.1%;   Pred. No. 1;
Matches 88;   Conservative 0;   Mismatches 81;   Indels 0;   Gaps 0;

Oy  88 AATAATTAAACAGTACGGGAATGCGAGAGCAATTTCTCTCTGGTGACAATATAGTGT 147
      ||| || || || || || || || || || || || || || || || || || || ||
Db  2479 ATTAGTGAACCTAGAACGAGAGACTCTGTGAGATAATCGATGTTTGTGTTTAAAGTTG 2538
      ||| || || || || || || || || || || || || || || || || || || ||

Oy  148 CCAACACTGCAAGTGATTTTAAAGATGTTTATTTAAATTAAGGATGATTTCCAAAG 207
      || || || || || || || || || || || || || || || || || || || ||
Db  2539 CTCAGTTTGGTCTAACTGTTGTATGCGACCAATAGATAAATAATATGCGAGAGAAAGAGAAA 2598
      ||| || || || || || || || || || || || || || || || || || || ||

Oy  208 GAAAAAATAAGCAAGGAAAGAAAAAACTGAACAGAAAAACGCAAAA 256
      |||| || || || || || || || || || || || || || || || || || || ||
Db  2599 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2647
      |||| || || || || || || || || || || || || || || || || || || ||

RESULT 15
US-09-149-476-66
; Sequence 66, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002p1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/JUS98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05

; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 2.8%; Score 39.2; DB 4; length 732;

Best Local Similarity 64.1%; Pred. No. 0.57;

Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 165 TTTTAAAGAATCTTTTATTTAAATTTAAAGGATGGATTTCACAGGAAAAAATAAGGAAA 224

Db 637 TTTATTAGGAGATTTTAAAGAAAAATAAGGTCTACAAATATCAAAAAAATAAAAAA 696

Qy 225 AGGAAGAAAAAACTGACACAGAAAAACGCAAAA 256

Db 697 AAAAAAATAAAAAAATAAAAAAATAAAAAA 728

Search completed: November 25, 2002, 18:45:38
Job time : 6027 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 16:00:01 ; Search time 2051 Seconds
(without alignments)
10912.805 Million cell updates/sec

Title: US-09-674-593-1
Perfect score: 1382
Sequence: 1 cattatgctaagcagcataaa.....taagcaattaagtctctgg 1382

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em.estba:*
2:	em.esthum:*
3:	em.estin:*
4:	em.estnu:*
5:	em.estov:*
6:	em.estpl:*
7:	em.estro:*
8:	em.htc:*
9:	gb.est1:*
10:	gb.est2:*
11:	gb.htc:*
12:	gb.est3:*
13:	gb.est4:*
14:	gb.est5:*
15:	em.estfun:*
16:	em.estom:*
17:	gb.gss:*
18:	em.gss_hum:*
19:	em.gss_inv:*
20:	em.gss_pln:*
21:	em.gss_vrt:*
22:	em.gss_fun:*
23:	em.gss_mam:*
24:	em.gss_mus:*
25:	em.gss_other:*
26:	em.gss_pro:*
27:	em.gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	776.2	56.2	879	14	BQ953400
2	730.2	52.8	1045	12	BG749180
3	665.6	48.2	845	14	BQ948660
4	609.6	44.1	898	14	BQ650550
5	534.4	38.7	963	14	BQ650598
6	470	34.0	486	14	BM833169

c 7	434.6	31.4	994	14	BQ650059
c 8	391.8	28.4	984	14	BM923267
c 9	288.4	20.9	929	14	BQ644671
c 10	288.4	20.9	939	14	BQ644474
c 11	288.4	20.9	1019	14	BM927835
c 12	252.8	18.3	477	14	BQ556967
c 13	252.8	18.3	634	10	BB660730
c 14	250.8	18.1	364	12	BF364659
c 15	244.4	17.7	246	9	AA863443
c 16	239.6	17.3	1048	14	BQ715572
c 17	232.8	16.8	923	12	BG751895
c 18	196.4	14.2	337	10	AW293816
c 19	192.2	13.9	835	12	BG750306
c 20	166.6	12.1	360	12	BF876266
c 21	145	10.5	683	17	AZ387429
c 22	122.8	8.9	525	13	BI841150
c 23	118.2	8.6	909	14	BQ731080
c 24	78.4	5.7	690	13	BJ072555
c 25	77.4	5.6	429	10	AW871951
c 26	75	5.4	665	14	BP027237
c 27	68.8	5.0	863	13	BI766821
c 28	67.8	4.9	635	10	BB663706
c 29	66.2	4.8	477	13	BI133726
c 30	66.2	4.8	541	14	BM933827
c 31	66.2	4.8	541	14	BM942392
c 32	66.2	4.8	806	13	BI916262
c 33	57.2	4.1	925	17	CNS0091P
c 34	57.2	4.1	925	17	CNS0091P
c 35	56	4.1	523	10	AW077096
c 36	55.4	4.0	925	17	CNS0091P
c 37	54.2	3.9	1101	17	CNS0039G
c 38	52.4	3.8	316	10	AW797167
c 39	52.4	3.8	1010	14	BQ668286
c 40	51.8	3.7	1125	9	AL547503
c 41	51.4	3.7	642	10	AV609566
c 42	51.2	3.7	319	9	AU257993
c 43	51.2	3.7	843	17	CNS01X1K
c 44	51	3.7	957	17	CNS015W7
c 45	50.8	3.7	354	14	BQ298642

ALIGNMENTS

RESULT 1
BQ953400
LOCUS
DEFINITION
AGENCOURT_8784199 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376296
5', mRNA sequence.
ACCESSION
BQ953400
VERSION
BQ953400.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 879)
NTH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: <http://image.lnl.gov>
Plate: LLCM2558 row: e column: 01
High quality sequence stop: 690.
FEATURES
source
1. .879

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6376296"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 194 a 275 c 236 g 173 t 1 others
ORIGIN

Query Match 56.2%; Score 776.2; DB 14; Length 879;
Best Local Similarity 98.8%; Pred. No. 2.7e-137;
Matches 792; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 524 AAATGGGTGGGGTGGGGAGACGACTCACTTGGATTCTTGAAGCTTCTCGGCCTC 583
Db 1 AAATGGGTGGGGTGGGGAGACGACTCACTTGGATTCTTGAAGCTTCTCGGCCTC 60
QY 584 CAGCCACGTAATTGCCCGCTCTGGATCTGGTCTAGCTTCGGATTTCGGTGGCCAGTCC 643
Db 61 CAGCCACGTAATTGCCCGCTCTGGATCTGGTCTAGCTTCGGATTTCGGTGGCCAGTCC 120
QY 644 GCGGGGTGTAGATGTCTCTAGCGGCCCAAGGGTGCCTGAACGCCGCGGTCACTCTCT 703
Db 121 GCGGGGTGTAGATGTCTCTAGCGGCCCAAGGGTGCCTGAACGCCGCGGTCACTCTCT 180
QY 704 TCAGGAAGACTTCGAAGCTGGACACTTCTTCATGTGATGACGACGCGCGCCCGGT 763
Db 181 TCAGGAAGACTTCGAAGCTGGACACTTCTTCATGTGATGACGACGCGCGCCCGGT 240
QY 764 AGAAGGGTCCCGTGGGTTACAAAGACGCTCTTCACGACGGGCTGAGACAGGTGGC 823
Db 241 AGAAGGGTCCCGTGGGTTACAAAGACGCTCTTCACGAC -GCCTGAGACAGGTGGC 299
QY 824 TGGACTTGGCGTGTGGCGCTATCTTCCCGCTGGCGCGCGCTCAGCTGCTGCTTC 883
Db 300 TGGACTTGGCGTGTGGCGCTATCTTCCCGCTGGCGCGCGCTCAGCTGCTGCTTC 359
QY 884 GGTGGGAGGACCTTCGCTGTCCAGGGCTTCACCGACCCAGGGCGGGATCGCC 943
Db 360 GGTGGGAGGACCTTCGCTGTCCAGGGCTTCACCGACCCAGGGCGGGATCGCC 419
QY 944 TCTGAAACGAACTGACGAATCCACAGGTGAAGAGAAAGTAAACGCCCGTGGC 1003
Db 420 TCTGAAACGAACTGACGAATCCACAGGTGAAGAGAAAGTAAACGCCCGTGGC 479
QY 1004 CTAGGGTCCACGAGAGAGACTAGGAGCTTGCAGGACTCGGAGTAGAGCTCAAGT 1063
Db 480 CTAGGGTCCACGAGAGAGACTAGGAGCTTGCAGGACTCGGAGTAGAGCTCAAGT 539
QY 1064 TTTTACCGTGGGTGCACAGCAATCAGACCCGAGTGGCGGACACACAGGTTCA 1123
Db 540 TTTTACCGTGGGTGCACAGCAATCAGACCCGAGTGGCGGACACACAGGTTCA 599
QY 1124 CTGTCTACGGGAGAAATCAAGGTGGACAGCTTCTGAGCAGGAGCGGAAACGCCGGG 1183
Db 600 CTGTCTACGGGAGAAATCAAGGTGGACAGCTTCTGAGCAGGAGCGGAAACGCCGGG 659
QY 1184 CTTCAACAGGACGCTTAGTAGGGCAGGAGAGGAGGAGGACACACACACACAC 1243
Db 660 CTTCAACAGGACGCTTAGTAGGGCAGGAGAGGAGGAGGACACACACACACAC 719
QY 1244 ACAAAATATGGTGAACCCCAATTTCTTACATCATATCTGTGTACCCCTTTTCCAAACAG 1303
Db 720 ACAAAATATGGTGAACCCCAATTTCTTACATCATATCTGTGTACCCCTTTTCCAAACAG 779

QY 1304 AAATTTCTTTTCTCTCTCTCTT 1325
Db 780 TAATTTTCTTTCTCTCTCTCTT 801
RESULT 2
BG749180 1045 bp mRNA linear EST 15-MAY-2001
LOCUS 602708139F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4844744 5',
DEFINITION mRNA sequence.
ACCESSION BG749180
VERSION BG749180.1 GI:14059833
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1682 row: n column: 09
High quality sequence stop: 822.
FEATURES
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4844744"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 245 a 322 c 300 g 178 t
ORIGIN
Query Match 52.8%; Score 730.2; DB 12; Length 1045;
Best Local Similarity 95.5%; Pred. No. 1.3e-128;
Matches 828; Conservative 0; Mismatches 28; Indels 11; Gaps 7;
QY 522 CTAATGGTGGGGTGGGGAGACCGACTCACTTGAAGCTTCTCGGC 581
Db 2 CTAATGGTGGGGTGGGGAGACCGACTCACTTGAAGCTTCTCGGC 61
QY 582 TCCAGCCAGTAAATGCCCGCTCTGGATCTGGATCTCCGATTCGGTGGCCAGT 641
Db 62 TCCAGCCAGTAAATGCCCGCTCTGGATCTGGATCTCCGATTCGGTGGCCAGT 121
QY 642 CCGCGGGGTAGATGTTCTTACGCCGCCCAAGGGTGCCTGAACGCCCGGTACCTC 701
Db 122 CCGCGGGGTAGATGTTCTTACGCCGCCCAAGGGTGCCTGAACGCCCGGTACCTC 181
QY 702 CTTAGGAAGACTTCGAAGCTGGACACTTCTTCTCATGGATGACACGCGGCGCCCGC 761
Db 182 CTTAGGAAGACTTCGAAGCTGGACACTTCTTCTCATGGATGACACGCGGCGCCCGC 241
QY 762 GTAGAAGGGGTCCCGTTCGGGTACACAGCAGCTCTTCACGACGGGCTGACAGAGT 821
Db 762 GTAGAAGGGGTCCCGTTCGGGTACACAGCAGCTCTTCACGACGGGCTGACAGAGT 821

Db	242	GTAGAAGGGTCCCCTTCCGGTACAAAGCACGCTCTTCAGCAC - GCCTGAGACAGGTG	300
Qy	822	GCTGGACCTGGCGCTGCTCCGCTCATCTTCCCCTGGCGCGCCCTCAGCTCGCTGCT	881
Db	301	GCTGGACCTGGCGCTGCTCCGCTCATCTTCCCCTGGCGCGCCCTCAGCTCGCTGCT	360
Qy	882	TCCGCTGGGAGGCACCTCCGCTGTCCAGCGGCTCACCGACCCAGGGCGCGGATCG	941
Db	361	TCGCGTGGGAGGCACCTCCGCTGTCCAGCGGCTCACCGACCCA - GGC GCGGGATCG	419
Qy	942	CCTCTGAAACGAACAGAGAACTGACGAATCCACAGGTGAAGAGAACTAACGCCGTGC	1001
Db	420	CCTCTGAAACGAACAGAGAACTGACGAATCCACAGGTGAAGAGAACTAACGCCGTGC	479
Qy	1002	GCCTAGGCGTCCACCCAGAGGAGACACTAGGAGCTTCGAGGACTCGGAGTAGAGCTCAA	1061
Db	480	GCCTAGGCGTCCACCCAGAGGAGACACTAGGAGCTTCGAGGACTCGGAGTAGAGCTCAA	539
Qy	1062	GTTCCTACGGTGGCGTGCA - CAGCCAACTCAGGACCCGAGTGC GCGCACACACAGGT	1120
Db	540	GTTCCTACCATGGCTGCAACAGCCAACTCAGGACCCGACGTGGCGGCACACACAGGT	599
Qy	1121	TCACCTGCTACGGGAGAAATCAAGTGTGACACTTCTGAGCAGGAGCCGGAACGCGCGG	1180
Db	600	TCACCTGCTACGGGAGAAATCAAGTGTGACACTTCTGAGCAGGAGCCGGAACGCGCGG	659
Qy	1181	GGCCTTCAACAGGCACGCTTAGTCAGGCGAGGAGAGGAGGCACACA - ---CACAC	1236
Db	660	GG - CTTCAACAGGCACGCTTAGTCAGGCGAGGAGAGGAGGCACACACAGAA	718
Qy	1237	CACACACACAAATATGGTGAACCCAAATTCATTACATCATATCTGTGTGCTACCCCTTCCAA	1296
Db	719	CACACACAAATATGGTGAACCCAAATTCATTACATCATATCTGTGTGCTACCCCTTCCAA	778
Qy	1297	ACAGCCTAATTTTCTTTTCTCTCTCTTGTGCACTTTTACCCCTC - AATCTCCTGCTTCCT	1355
Db	779	ACAGC - -TAATTTTCTTTTCTCTCTCTTGTGCACTTTTACCCCTCAAAATCTCCTGATGCT	836
Qy	1356	CCCAAAATTAAGCAATTAAGTTTCCTGG	1382
Db	837	CCCAAAATTAAGCAATTAAGTTTCCTGG	863

RESULT 3
LOCUS BQ948660/c
DEFINITION ACENCOURT_8784237 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376418
ACCESSION BQ948660
VERSION BQ948660.1 GI:22364138
KEYWORDS 5', mRNA sequence.
SOURCE EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.
COMMENT 1 (bases 1 to 845)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2558 row: j column: 03
High quality sequence stop: 534.
Location/Qualifiers
1..845
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:6376418"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC Library. |"
184 a 219 c 255 g 157 t 30 others
BASE COUNT
ORIGIN

```

Query Match	48.2%; Score 665.6; DB 14; Length 845;
Best Local Similarity	94.5%; Pred. No. 2.5e-116;
Matches	682; Conservative 0; Mismatches 39; Indels 1; Gaps
QY 245	GAAAGGCAAAAGATACAGTTTGTCACCTAACCTTTCGAAGGATACCTCTTTATTATTTCTT 304
DB 721	GAGNNNNNNNNNNNNNNNNNNNNNNNNNTNTTTCGAAGGATACCTTTTATTATTTCTT 662
QY 305	TAAAGATTCCGTGTTTATACACAGATTTTAAGTTTACTCCTACTGCTGCACCAGTGAA 364
DB 661	TAGGATCCCTGTTGTTTATACACAGA-TNTAAGTTTACTCCTACTGCTGCACCAGTGAA *603
QY 365	ATTCTCTCTCACGTCACAGTGTCAAACCTCTACCCCCAACCTGCACAGAGATTTTTGAGG 424
DB 602	ATTCTCTCTCAGTCACAGTGTCAAACCTCTACCCCCAACCTGCACAGAGATTTTTGAGG 543
QY 425	GCATCAATCACACCGAAGTCACAGCCCCTCAACCACTGAGTGTTGGGGGGTAGGGAT 484
DB 542	GCATCAATCACACGAGAGTCACAGCCCCTCAACCACTGAGTGTTGGGGGGTAGGGAT 483
QY 485	CTGCAATTTCTCATATAACCCCACTATATAGGCACTAAATGGGTGGCGGTGGGGGA 544
DB 482	CTGCATNTTCTCATATAACCCCACTATATAGGCACTAAATGGGTGGCGGTGGGGGA 423
QY 545	GACGCACTCACTTCAGTTTCTTGAAGCGTCTCTGGCCCTCCAGCCACGTAATTCGCCCCGC 604
DB 422	GACGCACTCACTTCAGTTTCTTGAAGCGTCTCTGGCCCTCCAGCCACGTAATTCGCCCCGC 363
QY 605	TCTGGATCTGCTAGCTTCCGGATTCCGTTGGCCAGTCCCGCGGGTGTAGATGTTCTGA 664
DB 362	TCTGGATCTGCTAGCTTCCGGATTCCGTTGGCCAGTCCCGCGGGTGTAGATGTTCTGA 303
QY 665	CGGCCCNAAAGGTCCTGAACGCCCGCGGTCACTCCTTTCAGAGACACTTCGAAGCTGG 724
DB 302	CGGCCCNAAAGGTCCTGAACGCCCGCGGTCACTCCTTTCAGAGACACTTCGAAGCTGG 243
QY 725	ACACTTCTTCTCATGGATGACAGCGCGCGCCCGGTAGAGAGGGGTCCCCCGTTGCGGT 784
DB 242	ACACTTCTTCTCATGGATGACAGCGCGCGCCCGGTAGAGAGGGGTCCCCCGTTGCGGT 183
QY 785	ACAAGACACGCTCTTCACGACGGGCTGAGACAGTGGCTGGACCTGGCGCTGCTGCCGC 844
DB 182	ACAAGACACGCTCTTCACGACGGGCTGAGACAGTGGCTGGACCTGGCGCTGCTGCCGC 123
QY 845	TCATCTTCCCGCTGGCGCGCTCAGCTCGCTTCGCGTCCGGAGGACCTCCGCT 904
DB 122	TCATCTTCCCGCTGGCGCGCTCAGCTCGCTTCGCGTCCGGAGGACCTCCGCT 63
QY 905	GTCCACGGCGCTCACCGCACCCAGGGCGGGGATCGCTCCTGAACAGAACGAGAACT 964
DB 62	GTCCACGGCGCTCACCGCACCCAGGGCGGGGATCGCTCCTGAACAGAACGAGAACT 3
QY 965	GA 966
DB 2	GA 1

FEATURES
SOURCE

```

RESULT 4
BQ650550/c
LOCUS
DEFINITION BQ650550 898 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8350342 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6286493
5', mRNA sequence.
ACCESSION BQ650550
VERSION BQ650550.1 GI:21774722
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2487 row: g column: 06
High quality sequence stop: 667.
Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6286493"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 206 a 252 c 273 g 167 t
ORIGIN
Query Match. 44.1%; Score 609.6; DB 14; Length 898;
Best Local Similarity 98.2%; Pred. No. 9.8e-106;
Matches 648; Conservative 0; Mismatches 9; Indels 3; Gaps 3;
QY 292 TTTTATTCTTTAAGATTCTCTGTT-GTTTATACAGATTTTAA-GTTTACTCTACT 349
DB 771 TTTTATTCTTTAAGGATTCCCGTTGGTTTATACAGATTTTAAAGGTTTACTCTACG 712
QY 350 GCTGACCAAGTGAATTCCTTCCAGTCACAGTGTCAACCTCTA-CCCCCAACTGCA 408
DB 711 GCTGACCAAGTGAATTCCTTCCAGTCACAGGTCACCTTACCCTCCCAACTGCA 652
QY 409 ACGAGAGTTTGGGGGCATCAATCACACCGAGAAGTCACAGCCCTCAACCACTGAGGT 468
DB 651 ACGAGAGTTTGGGGGCATCAATCACACCGAGAAGTCACAGCCCTCAACCACTGAGGT 592
QY 469 GTGGGGGGTAGGAGTGTGATTTCTTCATATCAACCCACACATATAGGCGACCTAAATG 528
DB 591 GTGGGGGGTAGGAGTGTGATTTCTTCATATCAACCCACACATATAGGCGACCTAAATG 532
QY 529 GTGGGGGGTAGGAGTGTGATTTCTTCATATCAACCCACACATATAGGCGACCTAAATG 588
DB 531 GTGGGGGGTAGGAGTGTGATTTCTTCATATCAACCCACACATATAGGCGACCTAAATG 472
QY 589 ACGTAATGCCCCCGTCTCGATCTGGTCTAGCTCCGGATTCGGTGGCCAGTCCCGGG 648
|||||

```

```

Db 471 ACGTAATGCCCCCGTCTCGATCTGGTCTAGCTTCGGATTCGGTGGCCAGCTCCGCGG 412
QY 649 GTGTAGATGTTCTGTAGGCGCCCAAGGGTGCCTGAACGCGCGGTCACTCTCTTCAGG 708
|||||
Db 411 GTGTAGATGTTCTGTAGGCGCCCAAGGGTGCCTGAACGCGCGGTCACTCTCTTCAGG 352
QY 709 AAGACTTCGAAGCTGGACACCTTCTTCTCATGATGACGACGCGGCGGCCGGGTAGGA 768
|||||
Db 351 AAGACTTCGAAGCTGGACACCTTCTTCTCATGATGACGACGCGGCGGCCGGGTAGGA 292
QY 769 GGTGCCCGTTGGGTATACAAAGACGCTCTTACAGAGGGGTGAGACAGGTGGCTGGAC 828
|||||
Db 291 GGTGCCCGTTGGGTATACAAAGACGCTCTTACAGAGGGGTGAGACAGGTGGCTGGAC 232
QY 829 CTGGCGCTGTCGGCTGCATCTTCCCGGTGGCGCGCTCAGCTCGCTTCGGCTC 888
|||||
Db 231 CTGGCGCTGTCGGCTGCATCTTCCCGGTGGCGCGCTCAGCTCGCTTCGGCTC 172
QY 889 GGGAGGCACCTCGCTGTCCAGCGGCTCACCGACCCAGCGGCGGGGATCGCTCTCTG 948
|||||
Db 171 GGGAGGCACCTCGCTGTCCAGCGGCTCACCGACCCAGCGGCGGGGATCGCTCTCTG 112
RESULT 5
BQ650598/c
LOCUS
DEFINITION BQ650598 963 bp mRNA linear EST 15-JUL-2002
AGENCOURT_8207577 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6283187
5', mRNA sequence.
ACCESSION BQ650598
VERSION BQ650598.1 GI:21774770
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2478 row: m column: 12
High quality sequence stop: 618.
Location/Qualifiers
1..963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6283187"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/notes="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 171 a 350 c 244 g 197 t 1 others
ORIGIN
Query Match 38.7%; Score 534.4; DB 14; Length 963;
Best Local Similarity 97.1%; Pred. No. 1.7e-91;
Matches 544; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

```
/clone="S5SNU484s1-12-C04"
/clone_lib="S5SNU484s1"
/sex="M"
/tissue_type="Stomach"
/cell_line="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/notes="Organ: Stomach; Vector: pTZ19Rpl; Site:1: EcoRI;
Site:2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dT-tailed
mRNA by priming with dT-tailed vector. The dT-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
After analyzing and sequencing about 2,000 ~ 3,000
colonies in original cDNA library, the abundant cDNAs were
selected and amplified by PCR reaction using vector region
primer including T7 promoter as 5' primer and N(drf)14 as
3' primer. The PCR products were used as template for
synthesis of biotinylated single stranded RNA by in vitro
transcription reaction. The synthesized RNA probes were
hybridized with antisense single stranded cDNAs prepared
from original library and incubated with avidin-gel.
After removing DNA-RNA hybrids by centrifuge, the
subtracted cDNA libraries were constructed by
transformation of the remaining DNA into competent cells E.
coli Top10F with electroporation method."
```

BASE COUNT 141 a 113 c 132 g 100 t
ORIGIN

Query Match 34.0%; Score 470; DB 14; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.3e-79;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 TTTGGTCACTAAACCTTTGCAAGGATACCTTTTATTTTCTTTAAGATTCCTGTGTTTAT 323
Db 470 TTTGGTCACTAAACCTTTGCAAGGATACCTTTTATTTTCTTTAAGATTCCTGTGTTTAT 411

Qy 324 ACACAGATTTTAAAGTCTACTGCTGACCCCAAGTGAATTCCTTCTCCAGTCACAG 383
Db 410 ACACAGATTTTAAAGTCTACTGCTGACCCCAAGTGAATTCCTTCTCCAGTCACAG 351

Qy 384 TGTCAACCTCTACCCCAACTGCAACGAGAGTTTGTAGGGGCGATCAATCAACCGAGAA 443
Db 350 TGTCAACCTCTACCCCAACTGCAACGAGAGTTTGTAGGGGCGATCAATCAACCGAGAA 291

Qy 444 GTCACAGCCCTCAACCACTCAGGTGTGGGGGTAGGATTCGATTCCTCATATCAA 503
Db 290 GTCACAGCCCTCAACCACTCAGGTGTGGGGGTAGGATTCGATTCCTCATATCAA 231

Qy 504 CCCACACTATAGGCGACCTAAATGGGTGGCGGTGGGGGAGACCGACTCACTTGAGTTT 563
Db 230 CCCACACTATAGGCGACCTAAATGGGTGGCGGTGGGGGAGACCGACTCACTTGAGTTT 171

Qy 564 CTTGAAGGCTTCTGGGCTCCAGCCAGTAATTTGCCCGCTCTGGATCTGGTCTAGCTT 623
Db 170 CTTGAAGGCTTCTGGGCTCCAGCCAGTAATTTGCCCGCTCTGGATCTGGTCTAGCTT 111

Qy 624 CCGGATTCGGTGGCGAGTCCCGGGGTGTAGATGTTCTTGACGCCGCCCAAGGGTGCCTG 683
Db 110 CCGGATTCGGTGGCGAGTCCCGGGGTGTAGATGTTCTTGACGCCGCCCAAGGGTGCCTG 51

Qy 684 AACGCCCGCGTCACTCTCTTCAGGAGACTTTCGAAGCTGGACACCTTCT 733
Db 50 AACGCCCGCGTCACTCTCTTCAGGAGACTTTCGAAGCTGGACACCTTCT 1

RESULT 7

389 ACCTCTACCCCACTGCAACGAGAGTTTGTAGGGGATCAATCAACCCAGAGTCCAC 448
Db 661 AACTGTCCCCCAATGCAAGGAGAGTTTGTAGGGGATCAATCAACCCAGAGTCC 602

Qy 449 AGCCCTCAACCACTGAGGTGTGGGGGTAGGATCTGCAATTTCTTCATATCAACCCCA 508
Db 601 AGCCCTCAACCACTGAGGTGTGGGGGTAGGATCTGCAATTTCTTCATATCAACCCCA 542

Qy 509 CACTATAGGCGACCTAAATGGGTGGGGGTGGGAGACCCACACCTTCTTGATTTCTTGA 568
Db 541 CACTATAGGCGACCTAAATGGGTGGGGGTGGGAGACCCACACCTTCTTGATTTCTTGA 482

Qy 569 AGGCTTCTGCGCTCCAGCCAGCTAAATTTGCCCGCTCTGATCTGCTAGCTTCCGGA 628
Db 481 AGGCTTCTGCGCTCCAGCCAGCTAAATTTGCCCGCTCTGATCTGCTAGCTTCCGGA 422

Qy 629 TTCGGTGGCCAGTCCGGGGGTGTAGATTTCTTCAGCGGCCCAAGGGTCCCTGAACGC 688
Db 421 TTCGGTGGCCAGTCCGGGGGTGTAGATTTCTTCAGCGGCCCAAGGGTCCCTGAACGC 362

Qy 689 CGCGGTCACTCTTTCAGGAGACTTTCGAAGCTGGACACCTTCTTCTCATGGATGACGA 748
Db 361 CGCGGTCACTCTTTCAGGAGACTTTCGAAGCTGGACACCTTCTTCTCATGGATGACGA 302

Qy 749 CGCGGTGGCCAGTCCGGGGGTCCCGGTTCGGGTACACAGCACGCTCTTCACGACGG 808
Db 301 CGCGGTGGCCAGTCCGGGGGTCCCGGTTCGGGTACACAGCACGCTCTTCACGACGG 242

Qy 809 GCTGAGACAGTGGCTGACCTGCGGTGTGCGGTGCTCATTTCCCGCTGGCGCGCCGCC 868
Db 241 GCTGAGACAGTGGCTGACCTGCGGTGTGCGGTGCTCATTTCCCGCTGGCGCGCCGCC 182

Qy 869 TCAGCTCGCTGCTGGGTTCGGGAGGACCTCCGCTGTCACGCGGCTCACCAGCCCA 928
Db 181 TCAGCTCGCTGCTGGGTTCGGGAGGACCTCCGCTGTCACGCGGCTCACCAGCCCA 122

Qy 929 GGGCGGGGATCGCTCCTG 948
Db 121 GGGCGGGGATCGCTCCTG 102

RESULT 6
BM833169/c
LOCUS
DEFINITION K-EST010772 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-12-C04
5', mRNA sequence.
ACCESSION BM833169
VERSION BM833169.1 GI:19189578
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 12 row: C column: 04
High quality sequence stop: 486.
location/Qualifiers
1. 486
/organism="Homo sapiens"
/db_xref="taxon:9606"

B0650059/c
LOCUS B0650059 994 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT 8297859 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269873
5', mRNA sequence.
ACCESSION B0650059
VERSION B0650059.1 GI:21774231
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 994)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2444 row: b column: 18
High quality sequence stop: 487.
FEATURES
source
1..994
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6269873"
/clone.lib="NIH_MGC_100"
/tissue.type="hepatocellular carcinoma, cell line"
/lab.host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 137 a 423 c 234 g 200 t
Query Match 31.4%; Score 434.6; DB 14; Length 994;
Best Local Similarity 94.8%; Pred. No. 1.3e-72;
Matches 460; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
QY 465 AGGTGTGGGGGGTGGGGATCTGCATTTCTTCATATCAACCCACACTATA-GGGCACCT 523
DB 585 AGTGATGAAGGTGAGATACGATGTGTGAAGCAGCCCGGCGCTAGGGCGCT 526
QY 524 AATAGTGGGGGGTGGGGAGACCGACTACTTGAAGTTCTTGAAGGCTTCTGGCCTC 583
DB 525 AATAGTGGAGCGTGGGGAGACCGACTTGAAGTTCTTGAAGGCTTCTGGCCTC 466
QY 584 CAGCCACGTAATGCCCCCGCTGTGATCTGTCTAGCTTCGGGATTCGGTGCGCCAGTCC 643
DB 465 CAGCCACGTAATGCCCCCGCTGTGATCTGTCTAGCTTCGGGATTCGGTGCGCCAGTCC 406
QY 644 GCGGGGTGTAGATGTTTCCTGAGCGGCCCAAGGGTGCCTGAACCGCGCGGTCACCTCTCT 703
DB 405 GCGGGGTGTAGATGTTTCCTGAGCGGCCCAAGGGTGCCTGAACCGCGCGGTCACCTCTCT 346
QY 704 TCAGGAAGACTTCGAAGCTGGACACCTTCTTCTCATGTGATGACGACGCGCGCGCGCT 763
DB 345 TCAGGAAGACTTCGAAGCTGGACACCTTCTTCTCATGTGATGACGACGCGCGCGCGCT 286
QY 764 AGAAGGGTCCCGCTTGGCGGTACACAAGCAGCCTCTTACGAGGGGCTGAGACAGGTGGC 823
DB 285 AGAAGGGTCCCGCTTGGCGGTACACAAGCAGCCTCTTACGAGGGGCTGAGACAGGTGGC 226

QY 824 TGGACCTGGCGCTGCTGCCGCTCATCTTCCCGCTGGCGCGCTCAGCTCGCTGCTTC 883
DB 225 TGGACCTGGCGCTGCTGCCGCTCATCTTCCCGCTGGCGCGCTCAGCTCGCTGCTTC 166
QY 884 GCGTCGGGAGGACCTCGGCTGTCCAGCGGCTCTACCGACCCAGGCGCGGGATCGCC 943
DB 165 GCGTCGGGAGGACCTCGGCTGTCCAGCGGCTCTACCGACCCAGGCGCGGGATCGCC 106
QY 944 TCCTG 948
DB 105 TCCTG 101
RESULT 8
LOCUS BM923267/c
DEFINITION AGENCOURT 6626059 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5758845
5', mRNA sequence.
ACCESSION BM923267
VERSION BM923267.1 GI:19373646
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 984)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12803 row: e column: 22
High quality sequence stop: 673.
FEATURES
source
1..984
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5758845"
/clone.lib="NIH_MGC_116"
/lab.host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT 246 a 254 c 279 g 205 t
Query Match 28.4%; Score 391.8; DB 14; Length 984;
Best Local Similarity 99.5%; Pred. No. 1.7e-64;
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 554 ACTTGAGTTCTTGAGGCTTCTGCGCTCCAGCCAGTAATTGCCCGCTCTGATCT 613
DB 517 AATTGAGTTCTTGAGGCTTCTGCGCTCCAGCCAGTAATTGCCCGCTCTGATCT 458
QY 614 GGTCTAGCTCCGGATTTCGGTGCGCCAGTCGCGGGGTGTAGATGTTCTTGACGCCCCCA 673
DB 457 GGTCTAGCTCCGGATTTCGGTGCGCCAGTCGCGGGGTGTAGATGTTCTTGACGCCCCCA 398

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 364)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-NN111-
070800-007-hl2&t3=2000-08-07&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 9.
Location/Qualifiers
1. .364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN111"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 82 a 114 c 112 g 56 t
ORIGIN

FEATURES

source

1. .364

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NN111"

/dev_stage="Adult"

/note="Organ: nervous_normal; Vector: puc18; Site.1: SmaI;

Site.2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 82 a 114 c 112 g 56 t

ORIGIN

Query Match 18.1%; Score 250.8; DB 12; Length 364;

Best Local Similarity 94.3%; Pred. No. 1.2e-37;

Matches 315; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

QY 860 GCGCGCGCTCAGCTCGCTTGGC-CTCGGAGGACCTCCGCTGTCAGGCGGCTC 918

DB 15 GCGTCGCGCTCAGCTCGCTTGGC-CTCGGAGGACCTCCGCTGTCAGGCGGCTC 74

QY 919 ACCGACCCAG-GGCGGGGATCGCTCTGAAAGCAAGCAACTGACGAATCCACAG 977

DB 75 ACCGACCCAGTGGCGGGATCGCTCTGAAAGCAAGCAACTGACGAATCCACAG 134

QY 978 GTGAAGAAGTAACGCGCGCTGCGCTAGGCGTCCACCCAGGAGACACTAGGAGCTT 1037

DB 135 GTGAAGAAGTAACGCGCGCTGCGCTAGGCGTCCACCCAGGAGACACTAGGAGCTT 194

QY 1038 GCAGACTCGGAGTAGAGCTCAAGTTTTTTCACCGTGGCGTGCA--CAGCCAAATCAGG-- 1093

DB 195 GCAGACTCGGAGTAGAGCTCAAGTTTTTTCACCGTGGCGTGCAACATGCCAATCAGGGA 254

QY 1094 ACCCGCAGTCGCGCAGCACACACAGGTTTCACCTGCTAGGGCAGAAATCAAGTGGAC-AG 1152

DB 255 CCGCGAGTTGCGCGCAGCACACACAGGTTTCACCTGCTAGGGCAGAAATCAAGTGGACAG 314

QY 1153 TTCTTGAGCAGGAGCGGAAACGCGGGGCGCTT 1186

DB 315 TTCTTGAGCAGGAGCGGAAACGCGGGGCGCTT 348

RESULT 15
AA863443

Search completed: November 26, 2002, 11:58:01

LOCUS

DEFINITION

AA863443

mRNA sequence.

AA863443

VERSION

AA863443.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 246)

NCI-CCGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

Tumor Gene Index

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-f@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image.html

Insert length: 1358 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 208.

Location/Qualifiers

1. .246

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1456938"

/clone_lib="NCI-CCGAP_Kid3"

/lab_host="DH10B"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st

strand cDNA was primed with a Not I - Oligo(dT) primer,

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT7T3 vector. mRNA

source: 2 pooled kidneys. Library went through one round

of normalization. Library constructed by Bento Soares and

M. Fatima Bonaldo."

BASE COUNT 63 a 66 c 44 g 73 t

ORIGIN

Query Match 17.7%; Score 244.4; DB 9; Length 246;

Best Local Similarity 99.6%; Pred. No. 2.2e-36;

Matches 245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 268 GTCACCTTTGCAAGGATACCTTTTATTTTCTTTAAGATTCTCTGTTTATACAC 327

DB 1 GTCATTACCTTTGCAAGGATACCTTTTATTTTCTTTAAGATTCTCTGTTTATACAC 60

QY 328 AGATTTTAAAGTTTACTCTCTACTGTCACCAAGTGAATTCCTTCACGTCACAGTGT 387

DB 61 AGATTTTAAAGTTTACTCTCTACTGTCACCAAGTGAATTCCTTCACGTCACAGTGT 120

QY 388 AACCTCTACCCGCCAAGTGAAGAGAGTTTGGGGGCATCAATCACCCGAGAAGTCA 447

DB 121 AACCTCTACCCGCCAAGTGAAGAGAGTTTGGGGGCATCAATCACCCGAGAAGTCA 180

QY 448 CAGCCCTCAACCTAGAGTGTGGGGGTAGGATTCGATTTCTCATATCAACCCC 507

DB 181 CAGCCCTCAACCTAGAGTGTGGGGGTAGGATTCGATTTCTCATATCAACCCC 240

QY 508 ACACCTA 513

DB 241 ACACCTA 246

Job time : 2077 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 08:44:00 ; Search time 265 Seconds
(without alignments)
11744.384 Million cell updates/sec

Title: US-09-674-593-1

Perfect score: 1382

Sequence: 1 cattatgctaacgcataaa.....taagcaattaagtctctg 1382

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1382	100.0	1382	21	AAZ36643 Human tumour rejec
C 2	595.4	43.1	2167	21	AAZ36644 Human tumour rejec
C 3	550	39.8	1956	22	AAI58192 Human polynucleoti
C 4	420.6	30.4	643	24	ABQ42610 Oligonucleotide fo
5	420.6	30.4	643	24	ABQ42611 Oligonucleotide fo
6	398.8	28.9	643	24	ABQ42612 Oligonucleotide fo
C 7	398.8	28.9	643	24	ABQ42613 Oligonucleotide fo
8	244.4	17.7	246	21	AAZ36649 EST AA863443 which
C 9	139.4	10.1	492	22	AAI59978 Human polynucleoti

C 10	68.4	4.9	5284	23	AA565026 DNA encoding novel
C 11	68	4.9	2424	22	AA444647 Novel protein kina
C 12	60	4.3	60	24	ABN38745 Human spliced tran
C 13	51.6	3.7	1645	21	AAF21831 Human breast and o
C 14	51.6	3.7	11735	22	AA543328 Chemically pretrea
C 15	51.6	3.7	11735	24	AA561141 Human gene regulat
C 16	51.6	3.7	11735	24	ABK28167 DNA transcrip
C 17	49.4	3.6	1937	23	ABV23182 Human prostate exp
C 18	48.4	3.5	823	22	ABA07664 Human ovarian and
C 19	48.4	3.5	823	22	AAI02457 Human reproductive
C 20	48.4	3.5	3509	21	AAZ57895 Arabidopsis RANL c
C 21	48.4	3.5	4800	21	AAZ57894 Arabidopsis RANL g
C 22	48	3.5	14987	24	ABL32630 Human immune syste
C 23	47.2	3.4	2482	24	ABK35668 cDNA sequence #59
C 24	47.2	3.4	513445	22	AAI61373 Soybean 318013 reg
C 25	46.4	3.4	353	22	AAI14789 Human breast cance
C 26	46	3.3	6092	24	AA561080 Human gene regulat
C 27	45.8	3.3	1754	13	AAQ20239 Gene encoding cell
C 28	45.4	3.3	5908	22	AA545386 Chemically pretrea
C 29	45.4	3.3	5908	24	AA561216 Human gene regulat
C 30	45.4	3.3	5908	24	ABK28231 DNA transcrip
C 31	45	3.3	2991	22	AAI15458 Human scavenger re
C 32	45	3.3	3019	22	AAI15459 Human scavenger re
C 33	45	3.3	3900	24	ABK15178 Human REPTR 10 CDN
C 34	44.8	3.2	1212	20	AAZ76499 Human WISP-3 prote
C 35	44.8	3.2	1212	20	AAZ76500 Human WISP-3 prote
C 36	44.6	3.2	1967	21	AAZ78171 Human cancer assoc
C 37	44.6	3.2	3265	22	AAZ08193 Human secreted pro
C 38	44.6	3.2	3265	22	AAZ05545 Human secreted pro
C 39	44.6	3.2	3339	22	AAZ05503 Human secreted pro
C 40	44.6	3.2	3339	22	AAZ05544 Human secreted pro
C 41	44.4	3.2	1159	22	AAZ02547 Human secreted pro
C 42	44.2	3.2	245	23	ABV07448 Human prostate exp
C 43	44.2	3.2	6106	22	AA546429 Tumour suppressor
C 44	44.2	3.2	6106	24	ABK40031 Human chemically p
C 45	44.2	3.2	6106	24	ABL33472 Human immune syste

ALIGNMENTS

RESULT 1

AAZ36643

ID AAZ36643 standard; cDNA; 1382 BP.

XX AC AAZ36643;

XX DT 22-FEB-2000 (first entry)

XX DE Human tumour rejection antigen RUR-1 antisense cDNA sequence.

XX KW Antisense; human; tumour rejection antigen; RUR-1; tumour; cancer;

XX KW renal cell carcinoma; colorectal carcinoma; melanoma; sarcoma;

XX KW leukaemia; ss.

XX OS Homo sapiens.

XX XX

XX Key Location/Qualifiers

XX CDS 738..992

XX FT /*tag= a

XX FT primer_bind 523..547

XX FT /*tag= b

XX FT /note= "binding site for primer VDE119 (see AAZ36647)"

XX FT complement (1280..1305)

XX FT /tag= c

XX FT /note= "binding site for primer VDE120 (see AAZ36648)"

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

FT CDS 303..1733
FT primer_bind /*tag= a
FT 484..501
FT /*tag= b
FT /note= "binding site for primer VDE87 (see AA236645)"
FT primer_bind complement (917..935)
FT /*tag= c
FT /note= "binding site for primer VDE93 (see AA236646)"
FT
FT
PN W0958546-A1.
XX
XX
PD 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US10424.
XX
XX 13-MAY-1998; 98US-0085318.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Van Den Eynde B, Boon-Falleur T;
XX
XX WPI: 2000-053076/04.
XX P-PSDB; AAY43811.
XX
XX New isolated tumour rejection antigen RUR-1 nucleic acids, used for,
XX e.g. treatment of cancers -
XX
XX Claim 5; Fig 6; 75pp; English.
XX
XX The present sequence represents the sense cDNA sequence of a human
XX ubiquitously expressed gene. The antisense strand of the present
XX sequence encodes tumour rejection antigen RUR-1. The antisense strand
XX codes for a polypeptide which is preferentially expressed in tumour
XX samples and tumour-derived cells lines. The polypeptide is unrelated
XX to any TRAP protein. The antisense sequence was isolated from a renal
XX cell carcinoma line L89211-RCC. The RUR-1 nucleic acids and polypeptides
XX can be used for diagnosis, prognosis or treatment of a disorder
XX characterized by the expression of a RUR-1 antisense cDNA molecule or
XX an expression product, such as cancers, e.g. renal cell carcinoma,
XX colorectal carcinoma, melanoma, sarcoma or leukaemia.
XX
XX Sequence 2167 BP; 675 A; 440 C; 576 G; 476 T; 0 other;
SQ

Query Match 43.1%; Score 595.4; DB 21; Length 2167;
Best Local Similarity 99.8%; Pred. No. 1.7e-135;
Matches 596; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 554 ACTTGAGTTCTTGAAGGCTTCCTGGCTCCAGCCAGTAAATGCCCGCTCTGGATCT 613
DB 597 AATTGAGTTCTTGAAGGCTTCCTGGCTCCAGCCAGTAAATGCCCGCTCTGGATCT 538

QY 614 GGCTAGCTCCGGATTCCGTGGCCAGTCCCGGGGTGTAGATGTTCTTGACGGCCCAA 673
DB 537 GGCTAGCTCCGGATTCCGTGGCCAGTCCCGGGGTGTAGATGTTCTTGACGGCCCAA 478

QY 674 AGGTGCTGTAACGCCCGCTCACCTCTTCAGGAAGACTTCGAAGCTGGACACTTCT 733
DB 477 AGGTGCTGTAACGCCCGCTCACCTCTTCAGGAAGACTTCGAAGCTGGACACTTCT 418

QY 734 TCTCATGGATGACGACGGCGGCCCGCGGTAGAAAGGGTCCCGGTTGCGGTACACAAGCA 793
DB 417 TCTCATGGATGACGACGGCGGCCCGCGGTAGAAAGGGTCCCGGTTGCGGTACACAAGCA 358

QY 794 CGCTCTTTCAGACGGGTGACAGAGTGGGTGGACCTGGCGCTGCTCCGCTCATCTTCC 853
DB 357 CGCTCTTTCAGACGGGTGACAGAGTGGGTGGACCTGGCGCTGCTCCGCTCATCTTCC 298

QY 854 CCGCTGCGCGCGCTCAGCTCGCTGCTCGCGTGGGAGGACCTCCGCTGTCACGAGCG 913
DB 297 CCGCTGCGCGCGCTCAGCTCGCTGCTCGCGTGGGAGGACCTCCGCTGTCACGAGCG 238

QY 914 GCCTCACCAGCCAGGCGCGGATCGCTCTCTTGAACGAACGAGAACTGACGAATCC 973
|||||

Db 237 GCCTCACCAGCCAGGCGCGGATCGCTCTTGAACGAACGAGAACTGACGAATCC 178
QY 974 ACAGGTGAAGAGAAAGTAACGGCGCTGCGCTAGGCGTCCACCCAGAGGAGACACTAGGA 1033
|||||
Db 177 ACAGGTGAAGAGAAAGTAACGGCGCTGCGCTAGGCGTCCACCCAGAGGAGACACTAGGA 118
QY 1034 GCTTGCAGGACTCGGAGTAGACGCTCAAGTTTTCACCGTGGCGTGCACAGCCCAATCAGG 1093
|||||
Db 117 GCTTGCAGGACTCGGAGTAGACGCTCAAGTTTTCACCGTGGCGTGCACAGCCCAATCAGG 58
QY 1094 ACCGCGAGTGGCGGCACACACACAGGTTCACTCTCTAGGCGCAGAACTCAAGGTGGAC 1150
|||||
Db 57 ACCGCGAGTGGCGGCACACACACAGGTTCACTCTCTAGGCGCAGAACTCAAGGTGGAC 1

RESULT 3
AA158192/c
ID AA158192 standard; cDNA; 1956 BP.
XX
XX AC AA158192;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 395.
DE
XX
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
XX
XX W0200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR
XX 09-JUL-2000; 2000US-0598042.
PR
XX 19-JUL-2000; 2000US-0620312.
PR
XX 03-AUG-2000; 2000US-0653450.
PR
XX 14-SEP-2000; 2000US-0662191.
PR
XX 19-OCT-2000; 2000US-0693036.
PR
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR
XX P-PSDB; AAM39036.
DR
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX
XX Claim 1; SEQ ID NO 395; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38442-AA42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX Homo sapiens.
OS
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX

DR WPI: 2001-442253/47.
DR P-PSDB; AAM40822.

XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

PS Claim 1; SEQ ID NO 3967; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 492 BP; 145 A; 111 C; 132 G; 100 T; 4 other;

Query Match 10.1%; Score 139.4; DB 22; Length 492;
Best Local Similarity 99.3%; Pred. NO. 3.2e-24;
Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 554 ACTTGAGTTCTTGAAGGTTCTCGGCTCCAGCCAGCAATGTCGCCGCTCTGGATCT 613

DB 193 AATTGAGTTCTTGAAGGTTCTCGGCTCCAGCCAGCAATGTCGCCGCTCTGGATCT 134

QY 614 GGTCTAGCTTCGGATTCGGTGGCAGTCCGCGGGGTGTAGATGTTCTGACGGCCCCAA 673

DB 133 GGTCTAGCTTCGGATTCGGTGGCAGTCCGCGGGGTGTAGATGTTCTGACGGCCCCAA 74

QY 674 AGGGTGCCTGAACGCCGCGG 694

DB 73 AGGGTGCCTGAACGCCGCGG 53

RESULT 10

AAS5026/c
ID AAS5026 standard; cDNA; 5284 BP.

XX AC AAS5026;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #830.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG00839.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity

PS Claim 1; SEQ ID NO 830; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS4197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5284 BP; 1370 A; 1419 C; 1627 G; 868 T; 0 other;

Query Match 4.9%; Score 68.4; DB 23; Length 5284;
Best Local Similarity 57.1%; Pred. NO. 2.1e-06;
Matches 145; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

QY 556 TTGAGTTCTTGAAGGTTCTCGGCTCCAGCCAGCAATGTCGCCGCTCTGGATCTGG 615

DB 4166 TGGAGCTTGTGGAATCGTTCAATCCAGCGCCACATCTGCCCTCTGTCTCAAGTCT 4107

QY 616 TCTAGCTTCCGGATTCGGTGGCCAGTCCGCGGGGTGTAGATGTTCTGACGGCCCCAAAG 675

DB 4106 GCCAGTTGGTGACAGGGTGGCCCATGCATGACAAAGTGTGTAGAGGCGCACACGCC ---ACT 4050

QY 676 GGTGCTGAACGCGCGGTACCTCTTCAAGAGACTTCAAGCTGGACACTTCTTC 735
 II IIIII I I IIIIIII IIIII I I I I I I
 Db 4049 GGGGCTGCACAGCTGATCTACCTCGCAGAGAGGCTCTCATGTGGGAGCGGCT 3990
 QY 736 TCATGATGACGACGCGCGCGGTAGAGGGTCCCCCTTCGGTACACAGCAGG 795
 I
 Db 3989 TGAGTCACACAGCTGGGAGCGCTGGGAAGATGGTCCCATTCGGGTACACCACTACC 3930
 QY 796 CTCTTCACGACGGG 809
 IIIII I I I I
 Db 3929 CTCTTGGCTGCTGG 3916
 RESULT 11
 ID AAF44647/c
 XX AAF44647 standard; cDNA; 2424 BP.
 AC AAF44647;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase cDNA, SEQ ID NO: 26.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000WO-US14842.
 PF
 PR 28-MAY-1999; 99US-0136503.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 DR WPI; 2001-032161/04.
 DR P-PSDB; AAB65621.
 XX
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers
 XX
 PS Example 1; Fig 2; 310pp; English.
 XX
 CC The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 2424 BP; 644 A; 575 C; 668 G; 535 T; 2 other;

Query Match 4.9%; Score 68; DB 22; Length 2424;
 Best Local Similarity 57.1%; Pred No. 1.9e-06;
 Matches 144; Conservative 0; Mismatches 105; Indels 3; Gaps 1;
 QY 558 GAGTTCTTGAAGGCTTCTCGCTCCAGCCACGTAATTTGCCCGCTCTGGATCTGGTC 617
 IIII
 Db 249 GAGCTTGTGGAATCGTTCAAATCCAGCGGCCACATACTGCCCTCTCTTCTTCAAGTCTGC 190
 QY 618 TAGCTTCCGGATTCTGGTGGCCAGTCCGGGGGTGTAGATGTTCTCTGACGCCCAAGG 677
 IIII
 Db 189 CAGGTGGTGACAGGTGGCCATGACAAGTGTGTAGAGGGCAGCCACGACCC---AGTGG 133
 QY 678 TGCCTGAACGCGCGGTCACCTCTTCAAGGAGACTTCAAGCTGGACACCTTCTTCTTC 737
 IIII
 Db 132 GCGCTGCACAGCTGATGTCACCTCGCAGAGGAGGCTCCATGGTGGGGAAGCGGTG 73
 QY 738 ATGGATGACGACGCGCGCGCGGTAGAGGGTCCCGCTGGCGGTAACAGACGACGCT 797
 I
 Db 72 AGTCACCAACGCTGGAGGCTGGGAAGATGGTCCCATTCGCGTACACCACTACCT 13
 QY 798 CTTCACGACGGG 809
 IIII I I I I
 Db 12 CTGGCTGCTGG 1
 RESULT 12
 ABN38745
 ID ABN38745 standard; DNA; 60 BP.
 XX
 AC ABN38745;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:11493.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-IB01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX
 DR WPI; 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 XX developmental-specific genes
 PS Example 1; SEQ ID 11493; 47pp; English.
 XX
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome. Where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal

CC transcriptsomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABB27253 to ABB59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 60 BP; 23 A; 15 C; 15 G; 7 T; 0 other;

Query Match 4.3%; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 938 ATCCCTCTCTGAAACGAGAACGAGAACTGACGAAATCCACAGGTGAAGAGAAAGTAACGGCC 997
 |||||
 Db 1 ATCCCTCTCTGAAACGAGAACGAGAACTGACGAAATCCACAGGTGAAGAGAAAGTAACGGCC 60

RESULT 13

AAF21831
 ID AAF21831 standard; DNA; 1645 BP.

AC AAF21831;

XX 27-MAR-2001 (first entry)

DE Human breast and ovarian cancer associated antigen gene SEQ ID 218.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW nototropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antibacterial; antifungal; antiparasitic; antitumor; anticonvulsant;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.

OS Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

DR P-PSDB; AAB58928.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -

XX Claim 1; Page 645-646; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the

CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antibacterial; antifungal; antiparasitic; antitumor; anticonvulsant;
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX

SQ Sequence 1645 BP; 400 A; 419 C; 448 G; 374 T; 4 other;

Query Match 3.7%; Score 51.6; DB 21; Length 1645;
 Best Local Similarity 60.9%; Pred. No. 0.016;
 Matches 84; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 119 AATTTTCTCCTCTGCTCACAATATAGTCTCCACACTTGGAAGTGATTTTAAAGAATGTT 178
 |||||
 Db 1469 AATTTTCTCCTCTGCTCACAATATAGTCTCCACACTTGGAAGTGATTTTAAAGAATGTT 1528

QY 179 TATTTAAATTAAGGATGGATTTCGAAGGAAAAAAATTAAGGAAGGAAAAAAGAAAAAC 238
 |||||
 Db 1529 TATTTCTCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1588

QY 239 TGAACAGAAACGCAAAA 256
 |||||

Db 1589 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1606

RESULT 14

AAS45328/c

ID AAS45328 standard; DNA; 11735 BP.

AC AAS45328;

XX 18-DEC-2001 (first entry)

DE Chemically pretreated genomic DNA associated with cell cycle #17.
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Levey body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.

OS Homo sapiens.

XX WO200168911-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02945.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

PI WPI; 2001-602751/68.

XX Designing primers and probes for analysing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,

THIS PAGE BLANK (USP10)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 05:31:33 : Search time 92.129 Seconds
(without alignments)
187.866 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDDAAPRVEGVPAVHKHA.....GAGSPPTNEKLTNPQVKEK 84

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	451	100.0	84	Q9UBP8	Q9ubp8 homo sapien
2	91.5	20.3	1154	Q9P2P0	Q9p2p0 homo sapien
3	91.5	20.3	1431	Q8WWD9	Q8wxd9 homo sapien
4	87	19.3	554	Q9X8S2	Q9x8s2 streptomyce
5	81	18.0	1083	O86637	O86637 streptomyce
6	80	17.7	456	Q8WU60	Q8wu60 homo sapien
7	80	17.7	456	Q96DP0	Q96dp0 homo sapien
8	80	17.7	527	Q96RE7	Q96re7 homo sapien
9	79.5	17.6	3247	Q65553	Q65553 bovine herp
10	78.5	17.4	438	O9C6J4	O9c6j4 arabidopsis
11	78	17.3	523	Q8RK12	Q8rk12 pseudomonas
12	78	17.3	678	Q8QZT8	Q8qzt8 mus musculu
13	77.5	17.2	1099	Q24096	Q24096 drosophila
14	77.5	17.2	1099	Q24590	Q24590 drosophila
15	77.5	17.2	1105	Q9VA38	Q9va38 drosophila
16	77	17.1	541	Q8RK09	Q8rk09 pseudomonas

17	77	17.1	700	5	P90929	P90929 caenorhabdi
18	77	17.1	1283	5	Q8T5H0	Q8t5h0 anopheles g
19	76.5	17.0	853	4	Q14333	Q14333 homo sapien
20	76	16.9	129	11	Q9CVD0	Q9cvd0 mus musculu
21	75.5	16.7	554	10	O04123	O04123 zea mays (m
22	75.5	16.7	843	10	Q94JY7	Q94jy7 arabidopsis
23	75.5	16.7	889	10	Q9FK48	Q9fk48 arabidopsis
24	75.5	16.7	1202	4	Q9P283	Q9p283 homo sapien
25	75.5	16.7	1413	5	Q9VJ38	Q9vj38 drosophila
26	75.5	16.7	1424	5	Q9VJ39	Q9vj39 drosophila
27	75	16.6	539	2	Q9RBW3	Q9rbw3 pseudomonas
28	75	16.6	1122	2	O85018	O85018 mycobacteri
29	75	16.6	1151	10	Q93VE9	Q93ve9 oryza sativ
30	75	16.6	1668	5	O76930	O76930 drosophila
31	74.5	16.5	298	16	O8UKS8	O8uks8 agrobacteri
32	74.5	16.5	3084	12	O8UZ11	O8uz11 pseudorabie
33	74	16.4	395	16	Q9RV74	Q9rv74 deinococcus
34	74	16.4	891	10	Q9ZW08	Q9zw08 arabidopsis
35	74	16.4	894	10	Q9FYB2	Q9fyb2 arabidopsis
36	73.5	16.3	251	10	Q9C6X7	Q9c6x7 arabidopsis
37	73.5	16.3	489	10	O9LNY9	O9lny9 arabidopsis
38	73.5	16.3	1740	10	O8S725	O8s725 oryza sativ
39	73	16.2	1171	5	Q9VHC2	Q9vhc2 drosophila
40	73	16.2	1198	2	O53730	O53730 streptomyce
41	72.5	16.1	118	11	Q9CVQ3	Q9cvq3 mus musculu
42	72.5	16.1	336	4	Q16096	Q16096 homo sapien
43	72.5	16.1	490	11	O8VCL9	O8vcl9 mus musculu
44	72.5	16.1	491	4	Q9H5S6	Q9h5s6 homo sapien
45	72.5	16.1	491	4	O8TAJ4	O8taj4 homo sapien

ALIGNMENTS

RESULT 1

Q9UBP8	PRELIMINARY;	PRT;	84	AA.
ID	Q9UBP8			
AC	Q9UBP8;			
DC	01-MAY-2000 (Tremblrel. 13, Created)			
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)			
DT	01-MAR-2001 (Tremblrel. 16, Last annotation update)			
DE	RU2AS protein.			
GN	RU2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20069887; PubMed=10601354;			
RA	Van den Eynde B.J., Gaugler B., Probst-Keppler M., Michaux L.,			
RA	Devuyt O., Lorge F., Weynants P., Boon T.;			
RT	"A new antigen recognized by cytolytic T lymphocytes on a human kidney			
RL	tumor results from reverse strand transcription.";			
RL	J. Exp. Med. 190:1793-1800(1999).			
DR	EMBL; AF181722; AAF23613.1;			
DR	ENBL; AF181720; AAF23611.1;			
SQ	SEQUENCE 84 AA; 8969 MW; 70B739F173A9E560 CRC64;			

Query Match 100.0%; Score 451; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5e-39;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDDDAAPRVEGVPAVHKHALHDGLROVAGFGAAGAAHLPRWPPPLQAASRRRAPPLSQRP	60
Db	1	MDDDAAPRVEGVPAVHKHALHDGLROVAGFGAAGAAHLPRWPPPLQAASRRRAPPLSQRP	60

Qy	61	HRTQAGSPPTNEKLTNPQVKEK	84
----	----	------------------------	----

Db	61	HRTQAGSPPTNEKLTNPQVKEK	84
----	----	------------------------	----

RESULT 2

```
Q9P2P0
ID Q9P2P0 PRELIMINARY; PRT; 1154 AA.
AC Q9P2P0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE KIAA1306 protein (Fragment).
GN KIAA1306.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037727; BAA92544.1; -.
DR HSPF; P29323; 184F.
DR InterPro; IPR002965; P_Rich_extensn.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00536; SAM; 2.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00454; SAM; 2.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
FT NON_TER 1
SQ SEQUENCE 1154 AA; 119815 MW; 05D603B5B98616B8 CRC64;

Query Match 20.3%; Score 91.5; DB 4; Length 1154;
Best Local Similarity 37.7%; Pred. No. 0.41;
Matches 23; Conservative 9; Mismatches 20; Indels 9; Gaps 2;

Qy 31 PGAAAHLPWPPLPPEGEARKPAKPVLPVTPVVKLGSPPTPTSKKVLPLPGSGSEV 981
Db 924 PPTDLAHLPLPPEGEARKPAKPVLPVTPVVKLGSPPTPTSKKVLPLPGSGSEV 983
Qy 82 K 82
Db 984 K 984

RESULT 3
Q8WXD9
ID Q8WXD9 PRELIMINARY; PRT; 1431 AA.
AC Q8WXD9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cask-interacting protein 1.
GN CASKIN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tabuchi K., Biederer T., Butz S., Sudhof T.C.;
RA "Cask participates in two alternative tripartite complexes in which
RT Mint 1 and Caskin 1 compete for binding.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF451977; AAL49758.1; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000104; Antifreeze-1.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR003993; treacle.

Q9X8S2
ID Q9X8S2 PRELIMINARY; PRT; 554 AA.
AC Q9X8S2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative spermidine synthase.
GN SC03655 OR SCH10.33C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
```

```
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00536; SAM; 2.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR00308; ANTIFREEZE1.
DR PRINTS; PR01503; TREACLE.
DR SMART; SM00248; ANK; 6.
DR SMART; SM00454; SAM; 2.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00088; ANK_REPEAT; 6.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS01005; SAM_DOMAIN; 2.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
SQ SEQUENCE 1431 AA; 149813 MW; CDF60E68B81E884A CRC64;

Query Match 20.3%; Score 91.5; DB 4; Length 1431;
Best Local Similarity 37.7%; Pred. No. 0.51;
Matches 23; Conservative 9; Mismatches 20; Indels 9; Gaps 2;

Qy 31 PGAAAHLPWPPLPPEGEARKPAKPVLPVTPVVKLGSPPTPTSKKVLPLPGSGSEV 1260
Db 1201 PPTDLAHLPLPPEGEARKPAKPVLPVTPVVKLGSPPTPTSKKVLPLPGSGSEV 1260
Qy 82 K 82
Db 1261 K 1261

RESULT 4
Q9X8S2
ID Q9X8S2 PRELIMINARY; PRT; 554 AA.
AC Q9X8S2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative spermidine synthase.
GN SC03655 OR SCH10.33C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
```

```

RT coellicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL049754; CAB42042.1;
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001045; Sprime_synthase.
DR Pfam: PF01564; Spermine_synth; 1.
SQ SEQUENCE 554 AA; 57758 MW; 593E484DCCD2DC42 CRC64;

Query Match 19.3%; Score 87; DB 16; Length 554;
Best Local Similarity 34.0%; Pred. No. 0.56;
Matches 35; Conservative 7; Mismatches 29; Indels 32; Gaps 7;

Qy 4 DAAPRVGVPVAVHKHALHD---GLRVAGVCAAAHLPRMPPO-----LAASRRAP- 54
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 457 DATLAAGLPAVAYRVGARDTGPVCGVGPVRAAG-P--PAPRGWGFLLAARTAPAPR 513
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 55 --PLSORPH-----RTQAGSPPETNEKLTNPQ 80
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 514 LDPAAPRPHSTPDSLARGVREARTOMAGLPPST---LVHPR 553
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
O86637 PRELIMINARY; PRT; 1083 AA.
AC O86637;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein SC05717.
GN SC05717 OR SC3C3.03C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL031231; CAA20252.1;
DR InterPro: IPR000194; ATPase_a/bcentre.
DR PROSITE: PS00152; ATPase_ALPHA_BETA; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1083 AA; 112084 MW; AD42EDFE4C12126D CRC64;

```

```

Query Match 18.0%; Score 81; DB 16; Length 1083;
Best Local Similarity 35.2%; Pred. No. 4.6;
Matches 25; Conservative 6; Mismatches 30; Indels 10; Gaps 2;

Qy 3 DDAAPRVGVPVAVHKHALHDGLRVAGVCAAAHLPRWP--PPOLAASRRAPLSORP 60
  ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 397 DDAAPPADSVP-----QDAVQDQATPPAPGPPPLPPSPVPAFAAAHQPAQPOPP 448
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 HRTQAGSGPPE 71
  : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 449 TOPQTAPAPPE 459

RESULT 6
O8WU60 PRELIMINARY; PRT; 456 AA.
ID O8WU60
AC O8WU60;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 49.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021217; AAH21217.1;
KW Hypothetical protein.
SQ SEQUENCE 456 AA; 49244 MW; E75515105D35269C CRC34;

Query Match 17.7%; Score 80; DB 4; Length 456;
Best Local Similarity 30.4%; Pred. No. 2.4;
Matches 24; Conservative 13; Mismatches 36; Indels 6; Gaps 3;

Qy 5 AAPRVGVPVAVHKHALHDGLRVAGVCAAAHLPR-WPPOLAASRRAPLSOR-PhR 62
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 284 APPKGPAPAVLAEEARAEGLQASPPGAAGHLEKVPPEPQ-----RRAAPLPKRKTS 339
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 63 TQAGSPPETNEKLTNPQ 81
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 340 TQTEGHLKGGQAIIVEDQI 358
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
O96DP0 PRELIMINARY; PRT; 456 AA.
ID O96DP0
AC O96DP0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDNA FLJ31479 fis, clone NT2NE2001634, moderately similar to
DE NADH-ubiquinone oxidoreductase 9 kDa subunit precursor
DE (EC 1.6.5.3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nimolaya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isodai T.;
RT "NEDO human cDNA sequencing project.";

```

```
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056041; BAB71080.1; -. C63A86DE137CDD59 CRC64;
SQ SEQUENCE 456 AA; 49245 MW; 63A86DE137CDD59 CRC64;

Query Match 17.7%; Score 80; DB 4; Length 456;
Best Local Similarity 30.4%; Pred. No. 2.4;
Matches 24; Conservative 13; Mismatches 36; Indels 6; Gaps 3;

QY 5 AAPRVEGVPAVVKHALHDLGRLQVAGPGAAAHLPRL-WPPQLAASRRREAPPLSQ-R-PHR 62
Db 284 APKSPAPAVIAEARAEGQLQASPPGAAGHLEKVPPEPQ-----RKAAPPLPRKETS 339
QY 63 TOGAGSPPETNEKLTNPQV 81
Db 340 TQIEGHLKGGQAIVEDQI 358

RESULT 8
Q96RE7 ID Q96RE7 PRELIMINARY; PRT; 527 AA.
AC Q96RE7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE NAC1 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Cha X.Y., Fakhrazadeh S.S.;
RT "Human NAC1 protein, a gene transcriptional repressor.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395817; AK83885.1; -.
DR InterPro; IPR000210; BTF_POZ.
DR Pfam; PF00651; BTF; 1.
DR PROSITE; PS50097; BTF; 1.
SQ SEQUENCE 527 AA; 57258 MW; 00BEA89D2C3DAD2 CRC64;

Query Match 17.7%; Score 80; DB 4; Length 527;
Best Local Similarity 28.3%; Pred. No. 2.8;
Matches 26; Conservative 13; Mismatches 29; Indels 24; Gaps 4;

QY 9 VEGVPVAVHKHALHDLGRLQVAGPGAAAHLPRLWPPPPQLAASRRREAPPLSQ-----RP 60
Db 176 VQCMPPVA--KRLWDSQKQAGGGGSGRKMKESTFDLAANRPHQPPPPQQAQPVVAAAQ 233
QY 61 HRTQAGSP-----PETNEKLTNP 79
Db 234 AVAAGAGQAGGVAAAGGVVSGPSTSER-TSP 264

RESULT 9
Q65553 ID Q65553 PRELIMINARY; PRT; 3247 AA.
AC Q65553;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE UL36.
GN UL36.
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RA Schwyzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.,
RA Thiry E., Paces V.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
```

RESULT 12
Q8QZT8
ID Q8QZT8 PRELIMINARY; PRT; 678 AA.

Query Match 17.2%; Score 77.5; DB 5; Length 1099;
Best Local Similarity 27.8%; Pred. No. 11;
Matches 27; Conservative 5; Mismatches 32; Indels 33

```

Qy 17 HKHALH-----DGLRQVAGP-----GAAAHLPWPVPQA 47
Db 162 HSHHTQSSRTVGNPGNGGFSFSEVAPPPRNPRTASSAATPPPPVPTQA 221
Qy 48 ASRREAPPLSQRPHR----TQAGSPPTNEKLTNPQ 80
Db 222 YVKRRSPALNNRPPAIPATPTQRGNSPVITONGLNKPQ 258

RESULT 14
Q24590 PRELIMINARY; PRT; 1099 AA.
AC Q24590;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Tumor suppressor.
GN WTS OR CG12072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95212904; PubMed=7698644;
RA Justice R.W., Zilian O., Woods D.F., Noll M., Bryant P.J.;
RT "The Drosophila tumor suppressor gene warts encodes a homolog of human
RT myotonic dystrophy kinase and is required for the control of cell
RT shape and proliferation.";
RL Genes Dev. 9:534-546(1995).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: L39837; AAF73959.1;
DR HSSP; P05132; 1APM.
DR FlyBase: FBgn0011739; wts.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; pkinase_C.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1099 AA; 121633 MW; EB35192A10470805 CRC64;

Query Match 17.2%; Score 77.5; DB 5; Length 1099;
Best Local Similarity 27.8%; Pred. No. 11;
Matches 27; Conservative 5; Mismatches 32; Indels 33; Gaps 3;

Qy 17 HKHALH-----DGLRQVAGP-----GAAAHLPWPVPQA 47
Db 162 HSHHTQSSRTVGNPGNGGFSFSEVAPPPRNPRTASSAATPPPPVPTQA 221
Qy 48 ASRREAPPLSQRPHR----TQAGSPPTNEKLTNPQ 80
Db 222 YVKRRSPALNNRPPAIPATPTQRGNSPVITONGLNKPQ 258

RESULT 15
Q9VA38 PRELIMINARY; PRT; 1105 AA.
AC Q9VA38;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE WTS protein.
GN WTS OR CG12072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```

```

OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottker P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003775; AAF57085.1;
DR HSSP; P05132; 1APM.
DR FlyBase: FBgn0011739; wts.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; pkinase_C.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1105 AA; 122488 MW; 7D09D7AA9BA1A856 CRC64;

Query Match 17.2%; Score 77.5; DB 5; Length 1105;
Best Local Similarity 27.8%; Pred. No. 11;
Matches 27; Conservative 5; Mismatches 32; Indels 33; Gaps 3;

Qy 17 HKHALH-----DGLRQVAGP-----GAAAHLPWPVPQA 47
Db 162 HSHHTQSSRTVGNPGNGGFSFSEVAPPPRNPRTASSAATPPPPVPTQA 221
Qy 48 ASRREAPPLSQRPHR----TQAGSPPTNEKLTNPQ 80

```


Db 222 YVRRSPALNNRPPAIAPTQRGNSPVITQNGLKNPQ 258

Search completed: November 26, 2002, 17:59:16
Job time : 107.129 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 11:58:11 ; Search time 46.0645 Seconds
(without alignments)
175.304 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDDAAPRVEGVVAVHKHA.....GAGSPETNEKLTNPQVKEK 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	19.3	554	2 T36545	probable spermidin
2	78.5	17.4	438	2 E96545	hypothetical prote
3	78	17.3	789	2 S44759	Cl4B9.5 protein -
4	77.5	17.2	1099	2 A56155	tumor suppressor p
5	77	17.1	482	1 S40887	RVS167 protein - y
6	77	17.1	700	2 T20350	hypothetical prote
7	75.5	16.7	554	1 T03263	calcium-dependent
8	75	16.6	1122	2 T14180	exit protein - Myc
9	75	16.6	1688	2 T13748	sex comb protein -
10	74.5	16.5	298	2 AD3162	conserved hypothet
11	74	16.4	395	2 H75457	hypothetical prote
12	74	16.4	891	2 G84693	probable proline-r
13	73.5	16.3	489	2 F96504	protein F9C16.29 [
14	73.5	16.3	817	2 T03852	protein phosphatas
15	73	16.2	1198	2 T28678	polyketide synthas
16	72.5	16.1	691	2 A25704	synapsin I - rat
17	72.5	16.1	704	2 A30411	synapsin Ia - rat
18	72.5	16.1	705	2 A35363	synapsin I splice
19	72.5	16.1	706	2 E30411	synapsin Ia - bovi
20	72	16.0	310	2 D70745	hypothetical prote
21	72	16.0	314	2 E84518	probable spliceoso
22	72	16.0	324	2 D82845	copper resistance
23	72	16.0	882	2 S41034	hypothetical prote
24	72	16.0	3164	1 WMBEH6	UL36 protein - hum
25	71.5	15.9	339	2 A43542	lymphocyte-specifi
26	71.5	15.9	428	2 S14426	dihydrolipoamide S
27	71.5	15.9	1324	2 T14070	peptide synthetase
28	71	15.7	563	2 T36580	hypothetical prote
29	70.5	15.6	490	2 A87672	tidd protein [impo

30	70.5	15.6	1026	2 T20369	hypothetical prote
31	70.5	15.6	2142	2 B35098	MHC class III hist
32	70	15.5	301	2 T36534	probable lipase/es
33	70	15.5	975	2 S33121	homeotic protein C
34	70	15.5	978	2 C59237	pyruvate dehydroge
35	70	15.5	1274	2 T37193	enamelin matrix pr
36	70	15.5	1298	1 EDBE75	immediate-early pr
37	70	15.5	1400	2 T31555	hypothetical prote
38	69.5	15.4	366	2 T26449	hypothetical prote
39	69.5	15.4	700	1 S01991	transforming prote
40	69.5	15.4	979	2 A70848	probable membrane
41	69	15.3	569	2 T19128	hypothetical prote
42	69	15.3	603	2 A75373	probable N-acetylm
43	69	15.3	865	2 T40170	hypothetical prote
44	69	15.3	1151	2 T18535	high molecular mas
45	68.5	15.2	257	2 T10586	small nuclear ribo

ALIGNMENTS

RESULT 1

T36545

probable spermidine synthase - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T36545

R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A;Reference number: 221609

A;Accession: T36545

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-554 <MUR>

A;Cross-references: EMBL:AL049754; PIDN:CAB42042.1; GSPDB:GN00070; SCOEDB:SCH10.33c

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SCH10.33c

Query Match 19.3%; Score 87; DB 2; Length 554;

Best Local Similarity 34.0%; Pred. No. 0.67;

Matches 35; Conservative 7; Mismatches 29; Indels 32; Gaps 7;

QY 4 DAAPRVEGVVAVHKHALHD---GLRQVAGPGAAAHLPWPPO-----LAASRREAP- 54

DB 457 DATRAAGLPVAVYRGVARDTGPVGSGVPPRRRAAG-P--PAPKNGFLLAARTAPAPR 513

QY 55 --PLSRQPH-----RTQGAGSPPETNEKLTNPQ 80

DB 514 LDPAAPRPHSLTPDSLARGVREAAQTOMAGLPST---LVHPR 553

RESULT 2

E96545

hypothetical protein F8A12.27 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: E96545

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: E96545

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-438 <STO>

R;Yu, S.; Fiss, E.; Jacobs Jr., W.R.
J. Bacteriol. 180, 4676-4685, 1998
A:Title: Analysis of the exochellin locus in mycobacterium smegmatis: biosynthesis gen
A:Reference number: Z17898; MUID:98389687; PMID:9721311
A:Accession: T14163
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 534-1122 <YUS>
A:Cross-references: EMBL:AF027770; NID:g3560502; PID:g3560505; PIDN:AACB2548.1
C:Genetics:
A:Gene: exit
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter Rv0194; ATP-bindin

Query Match 16.6%; Score 75; DB 2; Length 1122;
Best Local Similarity 31.0%; Pred. No. 19;
Matches 22; Conservative 9; Mismatches 28; Indels 12; Gaps 3;

QY 15 AVKHKALHD-GLRQVAGP-----AAAHLPRWPPPOLAAASREAPLSQRPHRTGAGS 68
| | | | | : | : | | : | | : | | : | |
DB 479 AAHARGTHNGGHQSGPARRGPGAARH-----RASGORRHTPFSGRRPLRLPGSGD 532
| | | | | : | : | | : | | : | | : | |

QY 69 PPETNEKLTPN 79
| : | : | | |
DB 533 PMISDEPVTPP 543

RESULT 9
T13748
sex comb protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13748
R:Sincclair, D.A.R.; Milne, T.A.; Hodgson, J.W.; Shellard, J.; Salinas, C.A.; Kyba, M.
Development 125, 1207-1216, 1998
A:title: The Additional sex combs gene of *Drosophila* encodes a chromatin protein that
A:Reference number: Z17750; MUID:98146384; PMID:9477319
A:Accession: T13748
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1668 <SIN>
A:Cross-references: EMBL:AJ001164; NID:g3292938; PIDN:CAA04568.1; PID:g3292939
C:Genetics:
A:Cross-references: FlyBase:FBgn0000142
C:Function:
A:Description: Involved in repression of homeotic loci

Query Match 16.6%; Score 75; DB 2; Length 1668;
Best Local Similarity 35.0%; Pred. No. 29;
Matches 21; Conservative 4; Mismatches 11; Indels 24; Gaps 3;

QY 13 PVAVKHALHDGLRQVAGPAAAHLPWPFPPO-----LAASHPEAPPLSORP-HRT 63
| | | | | : | | | | : | | : | | : | |
DB 676 PVLHR-----PAHRPHLPFPQCHLRAPVAIPAKRRRPQLHRRPEHRT 720

RESULT 10
AD3162
conserved hypothetical protein Atu5021 [imported] - *Agrobacterium tumefaciens* (strain
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AD3162
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.B.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl
: Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AD3162
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-298 <KUR>

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 17:06:07; Search time 88.5161 Seconds
(without alignments)
126.452 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDAARPRVEGVAVHKHA.....GAGSPETNEKLTNPQVKEK 84

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	451	100.0	84	21	AAV53809
2	86.5	19.2	1352	23	AAE18212
3	82.5	18.3	128	22	AAO11577
4	82	18.2	201	22	ABG03489
5	82	18.2	806	22	ABG04954
6	81.5	18.1	408	20	AAV41111
7	81.5	18.1	430	20	AAV41110
8	81.5	18.1	430	21	AAV70785
9	81.5	18.1	430	22	AAU09901
10	81.5	18.1	430	22	AAE05517

Protein encoded by Human MOLA protein Human polypeptide Novel human diagno Novel human diagno Human TANGO 129 (T Human TANGO 129 (T Human tumour necro Human tumour necro Human tumour necro

11	81.5	18.1	430	22	AAE95627	Human protein sequ
12	81.5	18.1	878	22	AAE02358	Human tumour necro
13	80	17.7	479	22	AAO13891	Human polypeptide
14	78.5	17.4	1078	22	ABG22869	Novel human diagno
15	77.5	17.2	160	22	ABG13164	Novel human diagno
16	77.5	17.2	1099	17	AAW05177	Lats large tumour
17	77.5	17.2	1099	21	AAV70393	Drosophila melanog
18	77.5	17.2	1105	22	ABE58401	Drosophila melanog
19	77	17.1	212	22	AAO12755	Human polypeptide
20	77	17.1	526	22	ABG20645	Novel human diagno
21	76.5	17.0	853	22	ABG12054	Novel human diagno
22	76.5	17.0	1072	21	ABG04157	Novel human diagno
23	76	16.9	111	21	ABG26081	zea mays protein f
24	76	16.9	621	22	AAU33270	Novel human secret
25	76	16.9	624	22	AAW42014	Human polypeptide
26	76	16.9	625	22	ABG21889	Novel human diagno
27	76	16.9	655	22	ABG19250	Novel human diagno
28	75.5	16.7	72	22	ABG13874	Novel human diagno
29	75.5	16.7	288	22	ABG21768	Novel human diagno
30	75.5	16.7	467	22	ABE11341	Human type II proc
31	75.5	16.7	1413	22	ABE60857	Drosophila melanog
32	75.5	16.7	1424	22	ABE60854	Drosophila melanog
33	75	16.6	510	22	ABG22923	Novel human diagno
34	75	16.6	906	22	AAE92916	Human protein sequ
35	74.5	16.5	120	22	AAO13454	Human polypeptide
36	74.5	16.5	174	22	ABG23384	Novel human diagno
37	73.5	16.3	202	22	ABG29088	Novel human diagno
38	73.5	16.3	700	11	AAE03663	Human myb related
39	73	16.2	93	22	AAU39584	Propionibacterium
40	73	16.2	138	22	AAU41659	Propionibacterium
41	73	16.2	1171	22	ABE64387	Drosophila melanog
42	72.5	16.1	193	22	ABE31899	Peptide #4550 enco
43	72.5	16.1	193	22	ABE37137	Peptide #4643 enco
44	72.5	16.1	193	22	ABE22450	Protein #4449 enco
45	72.5	16.1	193	22	AAE57857	Human brain expres

ALIGNMENTS

RESULT 1
AAV53809
ID AAV53809 standard; Protein: 84 AA.
XX
AC AAV53809;
XX
DT 22-FEB-2000 (first entry)
XX
DE Protein encoded by tmour rejection antigen RUR-1 antisense cDNA.
XX
KW Antisense; human; tumour rejection antigen: RUR-1; tumour; cancer;
KW renal cell carcinoma; colorectal carcinoma; melanoma; sarcoma;
KW leukaemia.
XX
OS Homo sapiens.
OS
XX
PN WO9558546-A1.
PN
XX
PD 18-NOV-1999.
XX
PF 13-MAY-1999; 99WO-US10424.
XX
PR 13-MAY-1998; 98US-0085318.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Van Den Eynde B, Boon-Falleur T;
XX
DR WPI; 2000-053076/04.
DR N-PSDB; AAZ36643.
XX
PT New isolated tumour rejection antigen RUR-1 nucleic acids, used for,
e.g. treatment of cancers -

XX Claim 16; Fig 5; 75pp; English.

XX The present sequence is encoded by the antisense cDNA sequence of human

CC tumour rejection antigen RUR-1. This cDNA sequence is the antisense

CC strand of a ubiquitously expressed gene. The antisense strand codes for

CC a polypeptide which is preferentially expressed in tumour samples and

CC tumour-derived cell lines. The polypeptide is unrelated to any TRAP

CC protein. The cDNA sequence was isolated from a renal cell carcinoma line

CC LB9211-RC. The RUR-1 nucleic acids and polypeptides can be used for

CC diagnosis, prognosis or treatment of a disorder characterized by the

CC expression of a RUR-1 antisense cDNA molecule or an expression product,

CC such as cancers, e.g. renal cell carcinoma, colorectal carcinoma,

CC melanoma, sarcoma or leukaemia.

XX Sequence 84 AA;

SQ

Query Match 100.0%; Score 451; DB 21; Length 84;

Best Local Similarity 100.0%; Pred. No. 4.1e-41;

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MDDAARVEGVAVVHKHALHDGLRQVAGPGCAAAHLPWPPLQLAASRRAPPLSORP 60

Db 1 MDDAARVEGVAVVHKHALHDGLRQVAGPGCAAAHLPWPPLQLAASRRAPPLSORP 60

Oy 61 HRTQAGSPPETNEKLTNPQVKEK 84

Db 61 HRTQAGSPPETNEKLTNPQVKEK 84

RESULT 2

AAE18212

ID AAE18212 standard; Protein; 1352 AA.

XX AAE18212;

AC AAE18212;

XX 07-MAY-2002 (first entry)

XX Human MOL4 protein.

XX Secreted molecule; MOL4 protein; MOLX; cardiomyopathy; atherosclerosis;

KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;

KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;

KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;

KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;

KW HIV; human immuno deficiency virus; hepatitis; haemostatic disease; pain;

KW haematopoietic; thrombolytic; thrombocytopaenia; Alzheimer's disease;

KW Parkinson's disease; Huntington's disease; muscular disease; stress;

KW ocular disease; growth disorder; depression; epilepsy; contraceptive;

KW vulneryary; osteopathic; haemostatic; tranquiliser; antidepressant;

KW analgesic; vasotropic; hypotensive; gene therapy.

XX Homo sapiens.

OS

XX W0200206339-A2.

PN

XX 24-JAN-2002.

XX

XX 03-JUL-2001; 2001WO-US21249.

XX

XX 03-JUL-2000; 2000US-215854P.

PR

XX 03-JUL-2000; 2000US-215856P.

PR

XX 03-JUL-2000; 2000US-215902P.

PR

XX 07-JUL-2000; 2000US-216585P.

PR

XX 07-JUL-2000; 2000US-216586P.

PR

XX 07-JUL-2000; 2000US-216722P.

PR

XX 17-JUL-2000; 2000US-218622P.

PR

XX 17-JUL-2000; 2000US-218922P.

PR

XX 27-JUL-2000; 2000US-221285P.

PR

XX 14-FEB-2001; 2001US-268734P.

XX

XX (CURA-) CURAGEN CORP.

XX

PI Spaderna SK, Tchernev V, Liu X, Shenoy S, Spylek K, Zerhusen B;

PI Patturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;

PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shimkets RA;

PI Padigaru M;

XX WPI; 2002-155038/20.

DR N-PSDB; AAD28946.

XX

XX Nucleic acids encoding secreted polypeptides, designated MOLX

PT polypeptides, useful for treating a MOLX-associated disorder, e.g.

PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -

XX Claim 1; Page 50; 223pp; English.

XX

CC The patent discloses nucleic acid sequences encoding novel secreted

CC molecule (MOLX) polypeptides, designated MOLX polypeptides (i.e. a MOL

CC protein where X is an integer from 1 to 8). Sequences of the invention

CC are useful for treating or preventing a MOLX-associated disorder in

CC humans. They are useful for treating or preventing cardiomyopathy,

CC atherosclerosis and disorders related to cell signal processing and

CC metabolic pathway modulation. The MOLX antibodies are useful for

CC treating or preventing diabetes and disorders related to cell signal

CC processing and metabolic pathway modulation. MOLX sequences are useful

CC for the treatment or diagnosis of other MOLX-associated disorders, e.g.

CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,

CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral

CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune

CC diseases, immunodeficiency, HIV (human immuno deficiency virus), viral,

CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,

CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopaenia,

CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,

CC ocular disease, muscular diseases, growth disorders, loss of libido,

CC stress, depression, pain and epilepsy. They are useful for preventing

CC chemotherapy side effects and as contraceptives. Sequences of the

CC invention are also useful for gene therapy. The present sequence

CC is human semaphorin-like protein, MOL4.

XX

SQ Sequence 1352 AA;

Query Match 19.2%; Score 86.5; DB 23; Length 1352;

Best Local Similarity 39.7%; Pred. No. 1.3;

Matches 27; Conservative 3; Mismatches 35; Indels 3; Gaps 2;

Oy 5 AARVEGVVAVVHKHALHD--GLRQVAGPGAAHLPWPPLQLAASRRAPPLSORPHR 62

Db 120 AHPAAGSPSAAGAGPGDCTGALRAGSCAAAPFPDPHPAHLVSSRRSAPPGSREPRG 179

Oy 63 TQAGSGSP 70

Db 180 T-GHLHPP 186

RESULT 3

AAO11577

ID AAO11577 standard; Protein; 128 AA.

XX

XX AAO11577;

XX

XX 06-NOV-2001 (first entry)

XX

XX Human polypeptide SEQ ID NO 25469.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

OS

XX W0200164835-A2.

PN

XX 07-SEP-2001.

XX

PF 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI: 2001-514838/56.
XX N-PSDB: AAI91508.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 20; SEQ ID NO 25469; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 128 AA;
SQ
Query Match 18.3%; Score 82.5; DB 22; Length 128;
Best Local Similarity 41.3%; Pred. No. 0.27;
Matches 19; Conservative 4; Mismatches 18; Indels 5; Gaps 2;
QY 27 QVAGPGAAAHLPWP---PPOLAASREAPPPLSORPHRTGAGSP 69
DB 36 QARGFGGAQAW--AWPGGPREAFSGGRRPPPLGLHPHRRKGAGPP 79
RESULT 4
ABG03489
ID ABG03489 standard; Protein; 201 AA.
XX
XX AC ABG03489;
XX
XX 13-FEB-2002 (first entry)
DT
XX Novel human diagnostic protein #3480.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI: 2001-639362/73.
XX

DR N-PSDB; AAS67676.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID NO 33848; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 201 AA;
SQ
Query Match 18.2%; Score 82; DB 22; Length 201;
Best Local Similarity 32.0%; Pred. No. 0.51;
Matches 32; Conservative 8; Mismatches 38; Indels 22; Gaps 5;
QY 3 DDAAAPRVEGVPVAVHKH---ALHDGLRQVA---GP-----GAAAHLPWPPLA 47
DB 20 DAAAHMAHPVPSHLRKQRVEVH-GFCQVQPSYCGEDRGLADRGSTDEHNPGAAQPRAA 78
QY 48 ASREAPPLSORPHRTGAGSPPTNEKL-----TNPQV 81
DB 79 ALHAHPGVGSQLPAPAHXAGQPPPTPEQLPVSPAXSNPQV 118
RESULT 5
ABG04954
ID ABG04954 standard; Protein; 806 AA.
XX
XX AC ABG04954;
XX
XX 13-FEB-2002 (first entry)
DT
XX Novel human diagnostic protein #4945.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI

XX WPI: 2001-639362/73.
 DR N-PSDB; AAS69141.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 35313; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain-reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 806 AA;
 Query Match 18.2%; Score 82; DB 22; Length 806;
 Best Local Similarity 30.2%; Pred. No. 2.3;
 Matches 26; Conservative 6; Mismatches 26; Indels 28; Gaps 3;
 QY 5 AAPRGVGVAVVHKHALHDGLRQVAGPGAAAHLPWPQPQLAASREAP----- 54
 Db 110 AAPVPAAPV-----PVAAPVPAAPVPRFQSPNOTASTSDSPHGQLPHQTV 157
 QY 55 PLS-----QRPHRTQGAGSPPTNE 74
 Db 158 PISDSFHTGQSPHRTSDPSPTGTQ 183
 RESULT 6
 ID AAY41111
 XX
 AC AAY41111;
 XX
 DT 17-JAN-2000 (first entry)
 DE Human TANGO 129 (T129) mature protein.
 XX
 KW Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
 KW cellular process; immunological disorder; abnormal lymphoid development;
 KW thymic development; T-cell mediated immune response; humoral B cell;
 KW skeletal muscle disorder; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO9952924-Al.
 XX
 PD 21-OCT-1999.
 XX
 PF 08-APR-1999; 99WO-US07832.
 XX
 PR 09-APR-1998; 98US-0057951.
 XX

(MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
 Holtzman D;
 WPI: 1999-620368/53.
 DR N-PSDB; AAZ23121.
 XX
 PT New isolated tumor necrosis factor receptor member used to develop
 PT products for treating, e.g. immunological disorders or disorders of the
 PT skeletal muscle
 XX
 PS Claim 8; Page 114-115; 118pp; English.
 XX
 CC The invention provides an isolated human tumor necrosis factor (TNF)
 CC receptor member, T129 (also referred as TANGO 129). The T129 polypeptide
 CC can be expressed by standard recombinant methodology. The T129
 CC polypeptides are useful as modulating agents in regulating a variety of
 CC cellular processes. Agents or modulators which have a stimulatory or
 CC inhibitory effect on T129 activity (e.g. T129 gene expression) as
 CC identified by a screening assay can be administered to individuals to
 CC treat (prophylactically or therapeutically) disorders, e.g. an
 CC immunological disorder associated with aberrant T129 activity, disorders
 CC associated with abnormal lymphoid and/or thymic development, T-cell
 CC mediated immune response, T-cell dependent help for B cells, and abnormal
 CC humoral B cell activity, and possibly disorders of the skeletal muscle.
 CC The products can also be used for detection, diagnosis, drug screening
 CC and production of transgenic animals. The present sequence represents
 CC the T129 mature polypeptide.
 XX
 SQ Sequence 408 AA;
 Query Match 18.1%; Score 81.5; DB 20; Length 408;
 Best Local Similarity 28.2%; Pred. No. 1.2;
 Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;
 QY 7 PRVEGVP-VAVHKHALH--DGLRQVAGPGAAAHLPWP---PPQLAASREAPPLSOR 59
 Db 244 PAPPNVPHICPHRHHLHTVQGLASLSGPGCCSRCSOKKWPVLLSPEAVANTPTVPSLLPN 303
 QY 60 PHRTOGAGSPPTNEKLT 77
 Db 304 PTRVPKAGAKAGRQEIT 321
 RESULT 7
 ID AAY41110
 XX
 AC AAY41110;
 XX
 DT 17-JAN-2000 (first entry)
 DE Human TANGO 129 (T129) protein.
 XX
 KW Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
 KW cellular process; immunological disorder; abnormal lymphoid development;
 KW thymic development; T-cell mediated immune response; humoral B cell;
 KW skeletal muscle disorder; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO9952924-Al.
 XX
 PD 21-OCT-1999.
 XX
 PF 08-APR-1999; 99WO-US07832.
 XX
 PR 09-APR-1998; 98US-0057951.
 XX
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
 PI Holtzman D;
 XX

appetite; memory; stress; food additive.

Homo sapiens.

Key Location/Qualifiers

Peptide

1..25
/label= Signal_peptide
/note= "Predicted leader sequence for TNF related gene 12"

Protein

26..430
/label= Mature_human_TNF_related_gene_12

/note= "Tumour necrosis factor related gene 12"

Domain

26..164

/note= "Predicted extracellular domain"

Domain

48..71

/note= "Predicted Cysteine rich domain"

Domain

165..181

/note= "Predicted transmembrane domain"

Domain

182..430

/note= "Predicted intracellular domain"

XX

WO200181402-A1.

XX

01-NOV-2001.

XX

19-APR-2001; 2001WO-US12762.

XX

19-APR-2000; 2000US-198388P.

PR

18-APR-2001; 2001US-0836607.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

Ruben SM, Ni J;

XX

WPI: 2001-616772/71.

DR

N-PSDB; AAS15031.

XX

Nucleic acid molecules encoding human TNFR Related Gene 12 proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

Claim 11: Fig 1A-C; 272pp; English.

The invention describes isolated nucleic acid molecules encoding Human Tumour necrosis factor receptor related gene 12 (I) proteins. These are used: in gene therapy (antisense) and vaccines; to prevent, or treat diseases associated with inappropriate expression of (I). The nucleic acids and complementary sequences may be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and identify patients needing restorative therapy. The (I) proteins are useful antigens in antibody production e.g. for detecting (I) in samples, and in assays to identify modulators of the expression and activity of the (I) proteins. The anti-(I) antibodies and antagonists may be used to down regulate expression and activity of the proteins. The disorders include: immune/autoimmune diseases (e.g. anaemia and rheumatoid arthritis), neurological diseases (e.g. Alzheimer's disease and Parkinson's disease) cancers and hyperproliferative disorders (e.g. neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and thrombosis), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). The agonists, antagonists, antibodies, polypeptides and polynucleotides can also be used to promote wound healing, maintain organs before transplantation, support cell culture of primary tissues, induce mesodermal tissue to differentiate in embryos, modulate differentiation of embryonic stem cells, modulate mammalian characteristics (e.g. height and weight), modulate e.g. catabolism, anabolism, mental state, biorhythms, cardiac rhythms, reproductive potential, appetite, memory and stress. The polypeptides can also be used as an additive to increase or decrease storage capabilities and nutritional content of food. This sequence is the human tumour necrosis factor (TNF) related gene 12 described in the method of the invention.

SQ Sequence 430 AA;

Query Match 18.1%; Score 81.5; DB 22; Length 430;

Best Local Similarity 28.2%; Pred. No. 1.3;

Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;

Qy 7 PRVEGVP-VAVHKHALH--DGLRQVAGPGAAAHLPWP----PPQLAASREAPPLSOR 59

Db 266 PAPPNVPHICPHRHHLHTVQGLASLSPCCSCRSQKKPEVLLSPFAVAATTPVPSLLPN 325

Qy 60 PHRTQAGSPETNEKLT 77

Db 326 PTRVPKAGAKAGQGEIT 343

RESULT 10

AAE05517

ID AAE05517 standard; Protein: 430 AA.

XX AAE05517;

XX 24-SEP-2001 (first entry)

XX Human tumour necrosis factor receptor/osteoprotegerin-like protein #1.

XX Human; cytostatic; tumour necrosis factor receptor/osteoprotegerin;

XX TNFR/OPG-like protein; antisense inhibitor; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..25

XX Protein 26..430

XX Domain 1..162

XX Domain 163..183

XX /label= Signal_peptide

XX /label= Mature_TNFR/OPG-like_protein

XX /label= Extracellular_domain

XX /label= Transmembrane_domain

XX WO200144472-A1.

XX 21-JUN-2001.

XX 14-DEC-2000; 2000WO-US33858.

XX 16-DEC-1999; 99US-0172306.

XX (AMGE-) AMGEN INC.

XX Jing S, Welcher AA, Fox GM, Shu J, Boedigheimer MJ, Bennett BD;

XX WPI: 2001-451665/48.

XX N-PSDB; AAD10353.

XX New tumor necrosis factor receptor / osteoprotegerin-like

XX (TNFR/OPG-like) polypeptides useful for diagnosis and treatment of

XX associated disease -

XX Claim 14; Fig 3; 208pp; English.

XX The present sequence is tumour necrosis factor receptor/osteoprotegerin

XX (TNFR/OPG)-like protein from human. TNFR/OPG-like protein is

XX useful for the treatment, prevention or amelioration of a medical

XX condition in a mammal resulting from decreased levels of TNFR/OPG

XX protein. Nucleic acid sequences of the present invention are used to

XX map locations of the TNFR/OPG genes and related genes. They are also

XX used as antisense inhibitors of TNFR/OPG expression. TNFR/OPG-like

XX protein and cDNA sequences are also used in gene therapy.

XX Sequence 430 AA;

Query Match 18.1%; Score 81.5; DB 22; Length 430;

CC	of the present invention.
XX	
SQ	Sequence 430 AA;
	Query Match 18.1%; Score 81.5; DB 22; Length 430; Best Local Similarity 28.2%; Pred. No. 1.3; Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;
QY	7 PRVEGVP-VAVHKLALH-DGLROVAGPGAAAAHLPRWP-----PPQLAASRREAPPLSQSQR 59 : : : : : Db 266 PAPNPVRICPHRHLLHTVOGLASLSPGCCSRCQSKKPEVLLSPVAVAATTPVPSLLPN 325
QY	60 PHRTQGAGSPETNEKLT 77 : :
Db	326 PTRVPKAGAKAGQEIT 343
RESULT 12	
AAE02358	ID AAE02358 standard; Protein; 878 AA.
XX	AAE02358;
AC	XX
XX	XX
DT	24-SEP-2001 (first entry)
XX	Human tumour necrosis factor receptor/osteoprotegerin-like protein #2.
DE	Human; cytostatic; tumour necrosis factor receptor/osteoprotegerin;
KW	TNFR/OPG-like protein; antisense inhibitor; gene therapy.
XX	Homo sapiens.
OS	XX
XX	XX
FH	Key Location/Qualifiers
FT	Misc-difference 495 /note= "Encoded by TGA"
FT	Misc-difference 499 /note= "Encoded by TAG"
FT	Misc-difference 564 /note= "Encoded by TGA"
FT	Misc-difference 645 /note= "Encoded by TAA"
FT	Misc-difference 658 /note= "Encoded by TGA"
FT	Misc-difference 663 /note= "Encoded by TAA"
FT	Misc-difference 677 /note= "Encoded by TGA"
FT	Misc-difference 679 /note= "Encoded by TGA"
FT	Misc-difference 695 /note= "Encoded by TAG"
FT	Misc-difference 714 /note= "Encoded by TAG"
FT	Misc-difference 715 /note= "Encoded by TAG"
FT	Misc-difference 747 /note= "Encoded by TAA"
FT	Misc-difference 764 /note= "Encoded by TAG"
FT	Misc-difference 766 /note= "Encoded by TGA"
FT	Misc-difference 776 /note= "Encoded by TAG"
FT	Misc-difference 788 /note= "Encoded by TAG"
FT	Misc-difference 824 /note= "Encoded by TGA"
FT	Misc-difference 872 /note= "Encoded by TAA"
FT	Misc-difference 873 /note= "Encoded by TAA"
XX	WO200144472-A1.
PN	

XX PD 21-JUN-2001.
XX PF 14-DEC-2000; 2000WO-US33858.
XX PR 16-DEC-1999; 99US-0172306.
XX PA (AMGE-) AMGEN INC.
XX PI Jing S, Welcher AA, Fox GM, Shu J, Boedigheimer MJ, Bennett BD;
XX WPI; 2001-451665/48.
DR N-PSDB; AAD10353.
XX New tumor necrosis factor receptor / osteoprotegerin-like
PT (TNFr/OPG-like) polypeptides useful for diagnosis and treatment of
PT associated disease -
XX Disclosure; Fig 5; 208pp; English.
XX The present sequence is tumour necrosis factor receptor/osteoprotegerin
CC (TNFr/OPG)-like protein from human. TNFr/OPG-like protein is
CC useful for the treatment, prevention or amelioration of a medical
CC condition in a mammal resulting from decreased levels of TNFr/OPG
CC protein. Nucleic acid sequences of the present invention are used to
CC map locations of the TNFr/OPG genes and related genes. They are also
CC used as antisense inhibitors of TNFr/OPG expression. TNFr/OPG-like
CC protein and cDNA sequences are also used in gene therapy.
XX SQ Sequence 878 AA;
Query Match 18.1%; Score 81.5; DB 22; Length 878;
Best Local Similarity 28.2%; Pred. No. 2.9;
Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;
QY 7 PRVEGVP-VAVHKHALH--DGLROVAGPAAAHLPWP-----PPQLAASRRREAPPLSQSR 59
Db 330 PAPPNVPHICPHRHHLHTVGLASLSGPGCCSGCKKWKPEVLLSPEAVAATTPVPSLLPN 389
QY 60 PHRTQAGSPPTNEKLT 77
Db 390 PTRVPKAGAKAGROGEIT 407
RESULT 13
AA013891
ID AA013891 standard; Protein; 479 AA.
XX AC AA013891;
XX 06-NOV-2001 (first entry)
XX Human polypeptide SEQ ID NO 27783.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX Homo sapiens.
OS WO200164835-A2.
XX 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US04927.
XX 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
PI

XX WPI; 2001-514838/56.
DR N-PSDB; AAI93822.
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX Claim 20; SEQ ID NO 27783; 1399pp + Sequence Listing; English.
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 479 AA;
Query Match 17.7%; Score 80; DB 22; Length 479;
Best Local Similarity 30.4%; Pred. No. 2.2;
Matches 24; Conservative 13; Mismatches 36; Indels 6; Gaps 3;
QY 5 AAPRVGVPVAVHKHALHLDGLROVAGPAAAHLPWP-----PPQLAASRRREAPPLSQSR-PHR 62
Db 307 APPKSGPAPVLAEEARAEQQLQASPPCAABGHLEKVPVPEPQ-----RKAAPPLPRKETSG 362
QY 63 TQAGAGSPPTNEKLTNPQV 81
Db 363 TQIEGHLKGGQAIVEDQI 381
RESULT 14
ABG22869
ID ABG22869 standard; Protein; 1078 AA.
XX AC ABG22869;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #22860.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS87056.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 18:45:47 : Search time 27.0968 Seconds
(without alignments)
128.577 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDDAAPRVEGVPAVHKHA.....GAGSPETNEKLTNPQVKEK 84

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result. No.	Score	Query	Length	DB	ID	Description
1	82.5	18.3	430	1	TRLT_MACFA	Q9092 macaca fasc
2	81.5	18.1	430	1	TRLT_HUMAN	Q96944 homo sapien
3	78	17.3	1780	1	YK26_CAEEL	P34333 caenorhabdi
4	77	17.1	482	1	R167_YEAST	P39743 saccharomyc
5	75	16.6	1032	1	S24D_HUMAN	O94855 homo sapien
6	73.5	16.3	494	1	SOX9_CHICK	P48434 gallus gall
7	73.5	16.3	817	1	NEB2_RAT	O35274 rattus norv
8	72.5	16.1	339	1	LSPI_HUMAN	P33241 homo sapien
9	72.5	16.1	670	1	SYN1_MOUSE	O88935 mus musculu
10	72.5	16.1	704	1	SYN1_RAT	P09951 rattus norv
11	72.5	16.1	705	1	SYN1_HUMAN	P17600 homo sapien
12	72.5	16.1	706	1	SYN1_BOVIN	P17599 bos taurus
13	72.5	16.1	719	1	DEND_RAT	P50617 rattus norv
14	72	16.0	310	1	Y497_MYCTU	Q11162 mycobacteri
15	72	16.0	2158	1	MY9B_HUMAN	Q13459 homo sapien
16	72	16.0	3164	1	TEGU_HSV11	P10220 herpes simp
17	71.5	15.9	427	1	ODP2_BACST	P11951 bacillus st
18	71	15.7	861	1	PQ58_CAEEL	P34552 caenorhabdi
19	71	15.7	1443	1	SIJ2_HUMAN	Q15056 homo sapien
20	70.5	15.6	2142	1	BAT2_HUMAN	P48634 homo sapien
21	70	15.5	943	1	NFC1_HUMAN	O95644 homo sapien
22	70	15.5	975	1	CUT1_CANFA	P39881 canis famil
23	70	15.5	1274	1	ENAM_MOUSE	O55196 mus musculu
24	70	15.5	1298	1	ICP4_HSV11	P08392 herpes simp
25	69.5	15.4	415	1	SYN1_CANFA	O62732 canis famil
26	69.5	15.4	700	1	MYBB_HUMAN	P10244 homo sapien
27	69.5	15.4	979	1	Y064_MYCTU	O53609 mycobacteri
28	69.5	15.4	1319	1	MN1_HUMAN	Q10571 homo sapien
29	69	15.3	1509	1	GSRI_HUMAN	Q9nzm4 homo sapien
30	68.5	15.2	359	1	Y4UE_RHISN	Q53198 rhizobium s
31	68.5	15.2	509	1	SOX9_HUMAN	P48436 homo sapien
32	68.5	15.2	604	1	MAON_HUMAN	Q16798 homo sapien
33	68.5	15.2	1238	1	BVGS_BORBR	P26762 bordetella

34	68.5	15.2	1238	1	BVGS_BORPA	P40330 bordetella
35	68.5	15.2	1409	1	AEX3_CAEEL	O02626 caenorhabdi
36	68	15.1	628	1	V70K_TYMW	P10357 turnip yell
37	68	15.1	1026	1	STAU_DROME	P25159 drosophila
38	67.5	15.0	615	1	MUTL_ECOLI	P23367 escherichia
39	67	14.9	113	1	N12B_MEDSA	Q40339 medicago sa
40	67	14.9	296	1	PRP3_MOUSE	P05143 mus musculu
41	67	14.9	309	1	HXA4_CHICK	P17277 gallus gall
42	67	14.9	421	1	FXJ1_HUMAN	Q92949 homo sapien
43	67	14.9	615	1	LM06_HUMAN	Q43900 homo sapien
44	67	14.9	622	1	3BP1_HUMAN	Q9Y313 homo sapien
45	67	14.9	1647	1	SN24_HUMAN	P51532 homo sapien

ALIGNMENTS

```
RESULT 1
ID TRLT_MACFA STANDARD; PRT; 430 AA.
AC Q9N092; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 1:FRSF19L precursor
DE (Receptor expressed in lymphoid tissues).
GN TNFRSF19L OR RELT.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21458551; PubMed=11574149;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RA Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K., Kusuda J.;
RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
RT chromosomes.";
RL Gene 275:31-37(2001).
CC -!- FUNCTION: Mediates activation of NF-kappa-B (by similarity). May
CC play a role in T-cell activation.
CC -!- SUBUNIT: Associates with TRAF1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- SIMILARITY: CONTAINS 1 TNFR-CYS REPEAT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB046039; BAB01621.1;
CC InterPro: IPR001368; TNFR_c6.
CC PROSITE: PS00652; TNFR_NGFR_1; FALSE_NEG.
CC SMART: PS00050; TNFR_NGFR_2; FALSE_NEG.
CC Receptor; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 26
CC CHAIN 27 430
CC TUMOR NECROSIS FACTOR RECEPTOR
CC SUPERFAMILY MEMBER TNFRSF19L.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC TNFR-CYS.
CC BY SIMILARITY.
CC BY SIMILARITY.
CC N-LINKED (GLCNAC...) (POTENTIAL).
CC CARBOHYD 149
CC SEQUENCE 430 AA; 45850 MW; BA8DE92593E1E859 CKC54;
```

Query Match 18.3%; Score 82.5; DB 1; Length 430;
Best Local Similarity 28.2%; Pred. No. 0.95;
Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;

QY 7 PRVEGVP-VAVHKHALH--DGLRQVAGPGAAAHLPWP-----PPQLAASRRREAPLSOR 59
Db 266 PAPPNPVPHICPHRHHLHTVOGLASLGSCSRCSQKKWPEVLLSPPEAVAATTSPSFLPN 325
QY 60 PHTQAGSPPTNEKLT 77
Db 326 PTPVPKAGAKAGQGEIT 343

RESULT 2
TRLT_HUMAN
ID TRLT_HUMAN STANDARD; PRT; 430 AA.
AC Q969Z4; Q96JUI; Q9BUX7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member TNFRSF19L precursor
DE (Receptor expressed in lymphoid tissues).
GN TNFRSF19L OR RELT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS.
RC TISSUE=Lymphoma;
RA Sica G.L., Zhu G., Tamada K., Liu D., Ni J., Chen L.;
RA MEDLINE=21213541; PubMed=11313261;
RT "RELT", a new member of the tumor necrosis factor receptor superfamily,
RT is selectively expressed in hematopoietic tissues and activates
RT transcription factor NF-kappaB."
RL Blood 97:2702-2707(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Retinoblastoma;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugeno S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwavanagi T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Eye;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 121-430 FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates activation of NF-kappa-B. May play a role in T-
CC cell activation.
CC -!- SUBUNIT: Associates with TRAF1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Highest levels are in spleen, lymph node,
CC thymus, peripheral blood leukocytes, bone marrow and fetal liver.
CC Very low levels in skeletal muscle, testis and colon. Not detected
CC in brain, kidney and pancreas.
CC -!- SIMILARITY: CONTAINS 1 TNFR-CYS REPEAT.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to several
CC frameshifts.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF319553; AAK77356.1; -
DR EMBL; AK027899; BAB5441.1; -
DR EMBL; BC001812; AAH01812.1; -
DR EMBL; BC017279; AAH17279.1; -
DR EMBL; AK074128; BAB84954.1; ALT_FRAME.
DR Genew; HGNC:13764; TNFRSF19L.
DR PROSITE; PS0052; TNFR_NGFR_1; FALSE_NEG.
DR PROSITE; PS0050; TNFR_NGFR_2; FALSE_NEG.
KW Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 430 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER TNFRSF19L.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT TNFR-CYS.
FT BY SIMILARITY.
FT DISULFID 51 65
FT DISULFID 71 90 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 122 122 D -> S (IN REF. 4).
FT CONFLICT 187 187 K -> E (IN REF. 2).
FT CONFLICT 273 273 H -> R (IN REF. 2).
FT CONFLICT 379 380 DL -> TR (IN REF. 3; AAH01812).
SQ SEQUENCE 430 AA; 46092 MW; 4A5AB9AE32D36101 CRC64;
Query Match 18.1%; Score 81.5; DB 1; Length 430;
Best Local Similarity 28.2%; Pred. No. 1.2;
Matches 22; Conservative 11; Mismatches 38; Indels 7; Gaps 3;

QY 7 PRVEGVP-VAVHKHALH--DGLRQVAGPGAAAHLPWP-----PPQLAASRRREAPLSOR 59
Db 266 PAPPNPVPHICPHRHHLHTVOGLASLGSCSRCSQKKWPEVLLSPPEAVAATTSPSFLPN 325
QY 60 PHTQAGSPPTNEKLT 77
Db 326 PTPVPKAGAKAGQGEIT 343

RESULT 3
YKZ6_CAEEL
ID YKZ6_CAEEL STANDARD; PRT; 1780 AA.
AC P34333; P34332;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C14B9.6 in chromosome III.
GN C14B9.6/C14B9.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulcon L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Woldman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be
CC produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15188; AAA27949.3; -;
DR EMBL: L15188; AAA27950.1; -;
DR PIR: S44758; S44758.
DR WormPep: C14B9.6a; CE000077.
DR WormPep: C14B9.6b; CE29577.
DR InterPro: IPR001005; MYB_DNA_bind.
DR SMART: SM00395; SANT; 1. FALSE_NEG.
DR PROSITE: PS00037; MYB_1; FALSE_NEG.
DR PROSITE: PS00334; MYB_2; FALSE_NEG.
DR PROSITE: PS50090; MYB_3; 1.
KW Hypothetical protein; DNA-binding; Nuclear protein;
KW Alternative splicing.
FT DNA_BIND 436 486
FT VARSPPLIC 949 1018
FT YSAEFAITFNSLKAQAQAOYQMRNPOLIPASVQMPIGINTA
FT YHOKSLTPSETASATPAHRPRAATTVG -> EMLGTLIN
FT AKYQFPQPOAIOQNPLMNAQHOLIQOQOLAMOHAAQA
FT AGSAGTSSSSSTGTGTSSRS (IN ISOFORM A).
FT MISSING (IN ISOFORM A).
FT VARSPLIC 1019 1780
FT SEQUENCE 1780 AA; 198569 MW; 0364817C41ABAD3F CRC64;
FT MYB.
Query Match 17.3%; Score 78; DB 1; Length 1780;
Best Local Similarity 29.5%; Pred. No. 10;
Matches 26; Conservative 13; Mismatches 25; Indels 24; Gaps 4;
QY 9 VEGVPVAVHKHLDGLRQVAGCAAAHLPRWPPPPQOLASRRRAPPLSQRPHR-----T 63
DB 1307 MQGAP---HRIH-----QLAGPSSSRISQSGRVSPALSRSVAPTISGAAGRSITAGT 1357
QY 64 QGAGSPPETNEKL-----TNPOV 81
DB 1358 SGHSSTSDYNKELAERNNEVFKPTHPQL 1385
RESULT 4
ID R167_YEAST STANDARD; PRT; 482 AA.
AC P39743;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Reduced viability upon starvation protein 167.
GN RVS167 OR YDR388W OR D9509.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-X2180;
RX MEDLINE-93330299; PubMed-8336735;
RA Bauer F., Urdaci M., Aigle M., Crouzet M.;

RT "Alteration of a yeast SH3 protein leads to conditional viability
RT with defects in cytoskeletal and budding patterns.";
RL Mol. Cell. Biol. 13:5070-5084(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunkeler-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Stroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP ACTIN-BINDING.
RX MEDLINE-95236199; PubMed-7719850;
RX Amerg D.C., Basart E., Botstein D.;
RT "Defining protein interactions with yeast actin in vivo.";
RL Nat. Struct. Biol. 2:28-35(1995).
CC -1- FUNCTION: COMPONENT OF A CYTOSKELETAL REORGANIZATION IN
CC FOR THE FORMATION OF ENDOCYTIC VESICLES AT THE PLASMA MEMBRANE
CC LEVEL. COULD BE IMPLICATED IN CYTOSKELETAL REORGANIZATION IN
CC RESPONSE TO ENVIRONMENTAL STRESSES AND COULD ACT IN THE BUDDING
CC SITE SELECTION MECHANISM. BINDS TO ACTIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE AMPHIPHYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M92092; AAA35051.1; -;
DR EMBL: U32274; AAB64830.1; -;
DR PIR: S40887; S40887.
DR HSP: P19174; 2HSP.
DR SGD: S0002796; RVS167.
DR InterPro: IPR004148; BAR_dom.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF03114; BAR; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Cytoskeleton; SH3 domain; Coiled coil; Transmembrane; Actin-binding.
KW COILED COIL (POTENTIAL).
KW COILED COIL (POTENTIAL).
FT DOMAIN 31 64
FT DOMAIN 174 204
FT DOMAIN 292 427
FT TRANSMEM 344 367
FT DOMAIN 421 482
FT SEQUENCE 482 AA; 3F0AB53EBCC95A5B CRC64;
Query Match 17.1%; Score 77; DB 1; Length 482;
Best Local Similarity 34.1%; Pred. No. 3.4;
Matches 28; Conservative 8; Mismatches 34; Indels 12; Gaps 4;
QY 5 AAPRVEGVPVAVHKHLDGLRQVAG--PGAAAAHLPRWPPPPQOLASRRRAPPLS----- 57
DB 332 AAPSYAAQPAQYCTAAAVGTAAAGVTAAGAAAGAVCTYPOYANA---QSPPLTGLGFO 388
QY 58 QRPHTQAGSPPETNEKLTNP 79
DB 389 QSPQOQO--GPPPAYSNPLTSP 408
RESULT 5
ID S24D_HUMAN STANDARD; PRT; 1032 AA.
AC O94855;
DT 16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Protein transport protein Sec24D (SEC24-related protein D).
SEC24D OR KIAA0755.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=99087487; PubMed=9872452;
Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
Tanaka A., Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
DNA Res. 5:277-286(1998).
[2]
SEQUENCE FROM N.A.
TISSUE=Pancreas;
MEDLINE=99262157; PubMed=10329445;
Tang B.-L., Kausalya J., Low D.Y.H., Lock M.L., Hong W.;
"A family of mammalian proteins homologous to yeast Sec24p.";
Biochem. Biophys. Res. Commun. 258:679-684(1999).
- FUNCTION: COMPONENT OF THE COPII COAT, THAT COVERS ER-DERIVED
VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
THE GOLGI APPARATUS. COPII ACTS IN THE CYTOPLASM TO PROMOTE THE
TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS
FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.
- SUBUNIT: COPII IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
COMPLEX, THE SEC13/31 COMPLEX AND SAR1.
- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR (BY SIMILARITY).
- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED, WITH HIGHER AMOUNTS IN
PLACENTA, PANCREAS, HEART AND LIVER.
- SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
EMBL; AB018298; BAA34475.1; -
EMBL; AF130464; AAD28756.2; -
Genew; HGNC:10706; SEC24D.
Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
Multigene family.
DOMAIN 363
FT DOMAIN 10 293 ZINC FINGER-LIKE.
PRO-RICH.
SEQUENCE 1032 AA; 112999 MW; 1EAA80215EB979A7 CRC64;
Query Match 16.6%; Score 75; DB 1; Length 1032;
Best Local Similarity 35.0%; Pred. No. 11;
Matches 21; Conservative 10; Mismatches 13; Indels 16; Gaps 4;
QY 30 GPCAAAHLPRW-----PPOLAASRRAPPLSRPHRTQOGASPPETNEKLTNPQV 81
DB 180 GPCASPLPLWYRPDGLSGPPDPN---AOYQPPPL---PGQTILGAGYPPQ--QANSQPQM 231
RESULT 6
SOX9_CHICK STANDARD; PRT; 494 AA.
AC F48434;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Transcription factor SOX-9.
GN SOX9.
OS Gallus gallus (Chicken).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauaria; Aves; Neognathae; Galliformes; Phasianinae;
Gallus.
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
Uwanogho D., Rex M., Cartwright E.J., Pearl G., Scotting P.J.,
Sharpe P.J.;
Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL SUBCELLULAR LOCATION: Nuclear (Potential).
CC -! SIMILARITY: CONTAINS 1 HMG BOX.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
EMBL; U12533; AAB09663.1; -
HSP; P48436; LSX9.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 106 174 HMG BOX.
FT DOMAIN 215 220 POLY-SER.
FT DOMAIN 342 360 GLN/PRO-RICH.
FT DOMAIN 394 401 POLY-GLN.
FT SEQUENCE 494 AA; 54942 MW; AF41D483B5016349 CRC64;
Query Match 16.3%; Score 73.5; DB 1; Length 494;
Best Local Similarity 27.5%; Pred. No. 7.4;
Matches 25; Conservative 12; Mismatches 37; Indels 17; Gaps 3;
QY 11 GPVVA---VVKHALHDLGRQVAGPGCAAAHL-----PRWPPPPQ---LAASRREA 53
DB 307 GVPATHGQVITYSGTYGSISSASSPAGAGHAWMAKQQPQPQPAQHTLPSTEREQ 366
QY 54 PPLSRPHRTQOGASPPETNEKLTNPQVKEK 84
DB 367 GPAQQRPHIKTEQLSPSHNSEQQOHPEOQQQ 397
RESULT 7
NEB2_RAT STANDARD; PRT; 817 AA.
ID NEB2_RAT
AC O35274;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurabin-II (Neural tissue-specific F-actin binding protein II).
DE (Protein phosphatase 1 regulatory subunit 9B) (Spinophilin) (p130)
DE (Pp1bpl34).
GN PP1R9B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97420791; PubMed=9275233;
RT Allen P.B., Ouimet C.C., Greengard P.;
"Spinophilin, a novel protein phosphatase 1 binding protein localized
to dendritic spines.";
Proc. Natl. Acad. Sci. U.S.A. 94:9956-9961(1997).
[2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-24; 164-182; 199-242; 449-480;
RP 678-689 AND 776-791.
RC TISSUE=Brain;
RX MEDLINE=98123121; PubMed=9452470;
SA Satoh A., Nakanishi H., Obaishi H., Wada M., Takahashi K., Satoh K.,

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X96589; CAA65407.1; -
FT NON_TER 1
SQ SEQUENCE 719 AA; 77208 MW; 89E2ED094514EAFB CRC64;

Query Match 16.1%; Score 72.5; DB 1; Length 719;
Best Local Similarity 33.9%; Pred. No. 13;
Matches 21; Conservative 5; Mismatches 17; Indels 19; Gaps 3;

QY 30 GPGAAAHNL-----PRW--PPQLAASRREAPLSORPHRTGAGSPPTNEKL 76

DB 211 GPPEAHLLLRGAAGTAPRRWRPPPV-----APPSYEGPHRTLTGKRGPELSRP 264

QY 77 TN 78

DB 265 TS 266

RESULT 14

Y497_MYCTU STANDARD; PRT; 310 AA.
ID Y497_MYCTU
AC Q11162;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV0497.
GN RV0497 OR MT0517 OR MTCY20C9.23.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=11773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv.
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; 277162; CAB00923.1; -
DR EMBL; AE006952; AAK4740.1; -
DR TIGR; MT0517; -

KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT DOMAIN 33 39 POLY-ARG.
FT DOMAIN 197 202 POLY-ALA.
SQ SEQUENCE 310 AA; 33092 MW; 4954027F694DF5C2 CRC64;

Query Match 16.0%; Score 72; DB 1; Length 310;
Best Local Similarity 30.4%; Pred. No. 6.4;
Matches 28; Conservative 9; Mismatches 29; Indels 26; Gaps 6;

QY 4 DAAPRVEGVPVAVHKKHALHDGLRQVAGPCGAAAHNL-----PRWP--PPQLAASRREA--PPL 56

DB 84 EAAQSPAEPA-----EQVAEPTRTVYNSQPEPRWPKSPQ---DRRESGPPL 130

QY 57 SORP-----HRTQAGSPPTNEKLTPQVK 82

DB 131 SEYPRPLRHTSDRAPAGPPSGAEHMSDPVE 162

RESULT 15

MY9B_HUMAN STANDARD; PRT; 2158 AA.
ID MY9B_HUMAN
AC Q13459; Q9UHN0; Q75314; Q9NUJ2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin IxB (unconventional myosin-9b).
GN MYO9B OR MYR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Small intestine;
RX MEDLINE=97063843; PubMed=8907710;
RA Wirth J.A., Jensen K.A., Post P.L., Bement W.M., Moosker M.S.;
RT "Human myosin-IxB, an unconventional myosin with a chimerin-like
RT rho/rac GTPase-activating protein domain in its tail."
RL J. Cell Sci. 109:653-661(1996).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=20047919; PubMed=10580159;
RA Grewal P.K., Jones A.-M., Maconochie M., Lemmers R.J.F., Frants R.R.,
RA Hewitt J.E.;
RT "Cloning of the murine unconventional myosin gene Myo9b and
RT identification of alternative splicing."
RL Gene 240:389-398(1999).
RN [3]
RP SEQUENCE OF 1940-2158 FROM N.A. (LONG ISOFORM).
RX MEDLINE=98158729; PubMed=9490638;
RA Post P.L., Bokoch G.M., Moosker M.S.;
RT "Human myosin-IxB is a mechanoschemically active molecule and a GAP for
RT rho."
RL J. Cell Sci. 111:941-950(1998).
RN [4]
RP SEQUENCE OF 1828-2023 FROM N.A. (SHORT ISOFORM).
RC TISSUE=Placenta;
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kinata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC MAY BE INVOLVED IN THE REMODELING OF THE INTRACELLULAR CYTOSKELETON. BINDS
CC ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
CC AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC ALSO ACTS AS A GPASE ACTIVATING PROTEIN ON RHO.
CC -!- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
CC F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
CC LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
CC PERINUCLEAR REGION.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN PERIPHERAL BLOOD
CC LEUCOCYTES AND AT LOWER LEVELS, IN THYMUS, SPLEEN, TESTIS,
CC PROSTATE, OVARY, BRAIN, SMALL INTESTINE, AND LUNG.
CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 IQ DOMAINS.
CC -!- CAUTION: THE C-TERMINAL SEQUENCE FROM AA 1917 ONWARDS FROM REF.1
CC WAS PROBABLY A CHIMERA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U42391; AAC50402.1; ALT_SEQ.
DR EMBL; AF143684; AAF00119.1; -
DR EMBL; AF020267; AAC26597.1; -
DR EMBL; AK002201; BAA92132.1; ALT_INIT.
DR HSSP; P08799; 1MND.
DR Genew; HGNC:7609; MYO9B.
DR MIM; 602129; -
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR000159; RA_domain.
DR InterPro; IPR000198; RhoGAP.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 2.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00612; IQ; 4.
DR Pfam; PF00620; RhoGAP; 1.
DR Pfam; PF00788; RA; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 2.
DR SMART; SM00109; C1; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR SMART; SM00314; RA; 1.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50096; IQ; 3.
KW Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW Cytoskeleton; Coiled coil; GTPase activation; Phorbol-ester binding;
KW Zinc; Alternative splicing.
FT DOMAIN 1 939 HEAD OR MOTOR DOMAIN.
FT DOMAIN 940 1044 NECK OR REGULATORY DOMAIN.
FT DOMAIN 1045 2158 TAIL.
FT DOMAIN 1046 1071 COILED COIL (POTENTIAL).
FT DOMAIN 1880 1901 COILED COIL (POTENTIAL).
FT DOMAIN 1959 1989 COILED COIL (POTENTIAL).
FT DOMAIN 844 855 ACTIN-BINDING.
FT DOMAIN 957 977 IQ 1.
FT DOMAIN 979 1000 IQ 2.
FT DOMAIN 1001 1023 IQ 3.
FT DOMAIN 1024 1053 IQ 4.
FT DOMAIN 1633 1681 PHORBOL-ESTER AND DAG BINDING.

FT DOMAIN 1713 1862 RHO-GAP.
FT NP_BIND 239 246 ATP (POTENTIAL).
FT VARSPLIC 2022 2023 PP -> OY (IN SHORT ISOFORM).
FT VARSPLIC 2024 2158 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1937 1939 QVP -> KT (IN REF. 4).
FT CONFLICT 1947 1947 L -> P (IN REF. 4).
FT CONFLICT 2040 2045 TVAAP -> PMPPLH (IN REF. 3).
FT CONFLICT 2049 2049 P -> L (IN REF. 3).
FT CONFLICT 2067 2067 P -> S (IN REF. 3).
FT CONFLICT 2157 2158 NG -> MAESHS (IN REF. 3).
SQ SEQUENCE 2158 AA; 243556 MW; 4978F1D770F56D28 CRC64;

Query Match 16.0%; Score 72; DB 1; Length 2158;
Best Local Similarity 29.2%; Pred. No. 43;
Matches 21; Conservative 8; Mismatches 19; Indels 24; Gaps 4;

QY 31 PGAAAAHLPRWPPQLAAS---RREAPPLSORPHR-----TQGA-----G 67
| :||||| | : || || :|| :
Db 2074 PGGLPSHLPRWAPGAREAAAPVRRRPP-ARRPDQIHVSITPGADLPVQCALEPLEEDG 2132

QY 68 SPPETNEKLTNP 79
|| : :||
Db 2133 QPPGAKRRYSDP 2144

Search completed: November 26, 2002, 12:02:21
Job time : 40.0968 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 12:00:06 ; Search time 29.8064 seconds
(without alignments)
82.919 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDDAAPRVGVPAVHRKHA.....GAGSPETNEKLTNPQVKRK 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	17.2	1099	US-09-442-100-2	Sequence 2, Appli
2	72.5	16.1	696	US-08-906-865-4	Sequence 4, Appli
3	72.5	16.1	696	US-09-129-668-4	Sequence 4, Appli
4	70	15.5	1298	US-08-690-473-2	Sequence 2, Appli
5	70	15.5	1298	US-09-259-821A-2	Sequence 2, Appli
6	70	15.5	1298	US-08-843-659-2	Sequence 2, Appli
7	68.5	15.2	509	US-08-860-635A-21	Sequence 21, Appl
8	68.5	15.2	509	US-09-281-476-21	Sequence 21, Appl
9	68	15.1	354	US-09-608-917A-2	Sequence 2, Appli
10	67.5	15.0	615	US-08-676-444-44	Sequence 44, Appl
11	67	14.9	543	US-09-535-008-63	Sequence 63, Appl
12	67	14.9	577	US-09-535-008-61	Sequence 61, Appl
13	67	14.9	1646	US-09-535-008-67	Sequence 67, Appl
14	67	14.9	1647	US-09-535-008-2	Sequence 2, Appli
15	67	14.9	1649	US-09-535-008-75	Sequence 75, Appl
16	67	14.9	1650	US-09-535-008-71	Sequence 71, Appl
17	67	14.9	1678	US-09-535-008-69	Sequence 69, Appl
18	67	14.9	1679	US-09-535-008-65	Sequence 65, Appl
19	67	14.9	1681	US-09-535-008-77	Sequence 77, Appl
20	67	14.9	1682	US-09-535-008-73	Sequence 73, Appl
21	66.5	14.7	625	US-09-347-801-18	Sequence 18, Appl
22	66	14.6	516	US-09-291-170A-1	Sequence 1, Appli
23	66	14.6	516	US-09-724-884-1	Sequence 1, Appli
24	65.5	14.5	404	US-09-232-468A-8	Sequence 8, Appli
25	65.5	14.5	3025	US-09-232-468A-8	Patent No. 5223423
26	65	14.4	834	US-08-491-357-2	Sequence 2, Appli
27	65	14.4	834	US-08-968-633-2	Sequence 2, Appli

28	65	14.4	834	3	US-09-196-466-2	Sequence 2, Appli
29	65	14.4	834	5	PCT-US96-10823-2	Sequence 2, Appli
30	65	14.4	905	2	US-08-574-959A-9	Sequence 9, Appli
31	65	14.4	905	4	US-09-357-014-9	Sequence 9, Appli
32	65	14.4	1135	2	US-08-574-959A-7	Sequence 7, Appli
33	65	14.4	1135	4	US-09-357-014-7	Sequence 7, Appli
34	65	14.4	2414	1	US-08-227-536-2	Sequence 2, Appli
35	65	14.4	2414	5	PCT-US95-04682-2	Sequence 2, Appli
36	65	14.4	2860	2	US-08-826-267-2	Sequence 2, Appli
37	64.5	14.3	865	4	US-09-281-766-19	Sequence 19, Appl
38	64	14.2	265	4	US-09-199-637A-369	Sequence 369, App
39	64	14.2	340	4	US-09-360-779-2	Sequence 2, Appli
40	64	14.2	340	4	US-09-435-335-2	Sequence 2, Appli
41	64	14.2	405	4	US-09-413-574-2	Sequence 2, Appli
42	64	14.2	591	1	US-08-484-840-2	Sequence 2, Appli
43	64	14.2	591	1	US-08-483-094-2	Sequence 2, Appli
44	63.5	14.1	2293	4	US-09-368-590-2	Sequence 2, Appli
45	63	14.0	432	2	US-08-937-466-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-442-100-2
; Sequence 2, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiwei
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1099 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-442-100-2
Query Match 17.2%; Score 77.5; DB 4; Length 1099;
Best Local Similarity 27.8%; Pred. No. 1.5;

Matches 27; Conservative 5; Mismatches 32; Indels 33; Gaps 3;

Qy 17 HKHALH-----DGLROVAGP-----GAAAHLPWWPPPOLA 47
Db 162 HSHHTQSSRTVGNPGNGGFSFSEVAPPPRNPNTASSAATPPPPVPTTQA 221

Qy 48 ASRREAPPLSORPHR---TQAGSPPTNEKLTNPQ 80
Db 222 YVKRSPALNRPPIAIPRTORGNSPVITQNGLNKNPQ 258

RESULT 2

US-08-906-865-4
; Sequence 4, Application US/08906865

; Patent No. 6040168

; GENERAL INFORMATION:

; APPLICANT: Greengard, Paul

; APPLICANT: Porton, Barbara

; APPLICANT: Kao, Hung-Teh

; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David A. Jackson, Esq.

; STREET: 411 Hackensack Ave, Continental Plaza, 4th

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/906.865

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-202

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 696 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; DESCRIPTION: /desc = "Synapsin Ia"

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

US-08-906-865-4

Query Match 16.1%; Score 72.5; DB 3; Length 696;

Best Local Similarity 35.8%; Pred. No. 3.3;

Matches 19; Conservative 7; Mismatches 18; Indels 9; Gaps 2;

Qy 26 RQVAGPAAAHLPWWPPPOLA-ASRREAPPLSORPHRTQAGSPPTNEKLT 77
Db 450 QQPAGPPAQEPPPGGPPPGPQPGPQPGPPLQQR-----PPQGGQHLS 494

RESULT 3

US-09-129-668-4

; Sequence 4, Application US/09129668B

; Patent No. 6429010

; GENERAL INFORMATION:

Query Match 15.5%; Score 70; DB 2; Length 1298;

Best Local Similarity 34.2%; Pred. No. 13;

Matches 25; Conservative 5; Mismatches 29; Indels 14; Gaps 3;

Qy 2 DDDAAPRVGVPVAVHKHALHDGLRQVAGPG-----AAAAHLPRWPPPOLAASRRAPPLS 57

; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-202 CIP
; CURRENT APPLICATION NUMBER: US/09/129.668B
; EARLIER FILING DATE: 1998-08-05
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-129-668-4

Query Match 16.1%; Score 72.5; DB 4; Length 696;

Best Local Similarity 35.8%; Pred. No. 3.3;

Matches 19; Conservative 7; Mismatches 18; Indels 9; Gaps 2;

Qy 26 RQVAGPAAAHLPWWPPPOLA-ASRREAPPLSORPHRTQAGSPPTNEKLT 77
Db 450 QQPAGPPAQEPPPGGPPPGPQPGPQPGPPLQQR-----PPQGGQHLS 494

RESULT 4

US-08-690-473-2

; Sequence 2, Application US/08690473

; Patent No. 5876923

; GENERAL INFORMATION:

; APPLICANT: Leopardi, Rosario

; APPLICANT: Roizman, Bernard

; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN

; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/690.473

; FILING DATE: 26-JUL-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 37,642

; REFERENCE/DOCKET NUMBER: ARCD:239

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/418-3000

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1298 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-690-473-2

```

Db      552  DDDAGRRAGRAVAVECLAACRGILEALAEFGDILAAV-----PGLAGRAPASPP---602
QY      58   QRPRTQGAGSPP 70
Db      603  -RPEGPAGPASPP 614

```


RESULT 15
US-09-535-008-75
; Sequence 75, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtqian, Sean V.

Wed Nov 27 08:47:42 2002

APPLICANT: Teng, David H.-F.
TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
FILE REFERENCE: 2318-259
CURRENT APPLICATION NUMBER: US/09/535,008
CURRENT FILING DATE: 2000-03-23
EARLIER APPLICATION NUMBER: U.S. 60/125,806
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 75
LENGTH: 1649
TYPE: PRT
ORGANISM: Homo sapiens
US-09-535-008-75

Query Match 14.9%; Score 67; DB 4; Length 1649;
Best Local Similarity 27.4%; Pred. No. 37;
Matches 20; Conservative 6; Mismatches 31; Indels 16; Gaps 2;
Qy 11 GVPVAVHKHALHDGLROVAGPCAAAHLPRWPPOLAAS-----RREAPPLS 57
Db 193 GQPLPDHLQMAVQGRMPG---MOQQMTLPPPSVATGPGPGPGPGPPNY 249
Qy 58 QRPHTQAGAGSPP 70
Db 250 SRPHCMGGFNMP 262

Search completed: November 26, 2002, 19:43:18
Job time : 33.8064 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2002, 12:01:36 : Search time 44.2581 Seconds
(without alignments)
30.223 Million cell updates/sec

Title: US-09-674-593-2

Perfect score: 451

Sequence: 1 MDDDAAPRVEGVAVHKA.....GAGSPETNETKLTPQVKEK 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	81.5	18.1	408	12	US-09-057-951-4
2	81.5	18.1	408	10	US-10-105-150-4
3	81.5	18.1	430	10	US-09-057-951-2
4	81.5	18.1	430	10	US-09-836-607-2
5	81.5	18.1	430	12	US-10-105-150-2
6	72.5	16.1	193	10	US-09-864-761-37748
7	69.5	15.4	538	12	US-10-023-529-43
8	69.5	15.4	538	12	US-10-023-523-43
9	68.5	15.2	509	10	US-09-910-087-21
10	68	15.1	684	10	US-09-823-240-9
11	67	14.9	721	9	US-09-764-868-731
12	67	14.9	920	10	US-09-983-531A-4
13	66.5	14.7	625	10	US-09-854-731-18
14	66.5	14.7	846	10	US-09-858-664A-3
15	66.5	14.7	4019	10	US-09-738-973-425
16	65.5	14.5	531	10	US-09-925-300-1444
17	65	14.4	494	10	US-09-833-790-234
18	64.5	14.3	190	10	US-09-864-761-36101
19	64.5	14.3	419	10	US-09-814-777A-36

20	64.5	14.3	865	9	US-09-957-995A-19
21	64	14.2	340	10	US-09-850-799-2
22	64	14.2	405	10	US-09-805-550-2
23	64	14.2	430	12	US-10-062-254-252
24	64	14.2	888	12	US-10-036-342-35
25	63.5	14.1	104	10	US-09-864-761-33381
26	63.5	14.1	339	10	US-09-745-763-146
27	63.5	14.1	339	12	US-10-036-342-80
28	63.5	14.1	339	12	US-10-052-586-504
29	63.5	14.1	503	12	US-10-062-254-240
30	63	14.0	184	10	US-09-764-864-827
31	63	14.0	302	10	US-09-764-864-1331
32	63	14.0	520	10	US-09-726-643-102
33	63	14.0	526	10	US-09-925-300-1473
34	63	14.0	591	10	US-09-962-290-4
35	62.5	13.9	280	10	US-09-823-038A-18
36	62	13.7	507	10	US-09-910-087-19
37	62	13.7	559	10	US-09-877-633-12
38	62	13.7	596	12	US-10-006-867-100
39	62	13.7	596	12	US-10-052-586-310
40	61.5	13.6	283	12	US-10-052-586-194
41	61.5	13.6	321	9	US-09-945-182-26
42	61.5	13.6	480	10	US-09-790-264-32
43	61.5	13.6	756	9	US-09-992-598-62
44	61.5	13.6	756	10	US-09-989-722-62
45	61.5	13.6	756	10	US-09-989-723-62

ALIGNMENTS

RESULT 1
US-09-057-951-4
; Sequence 4, Application US/09057951
; Patent No. US2002025551A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; TITLE OF INVENTION: NOVEL MOLECULES OF THE T129-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,951
; FILING DATE: 09-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 09404/046001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-057-951-4
Query Match 18.1%; Score 81.5; DB 10; Length 408;
Best Local Similarity 28.2%; Pred. No. 0.36;


```

; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-523-43

Query Match          15.4%; Score 69.5; DB 12; Length 538;
Best Local Similarity 37.3%; Pred. No. 8.2;
Matches 25; Conservative 2; Mismatches 31; Indels 9; Gaps

QY 5 AARPEGVPAVVKHHLHDGLRQVAGCAAAHLPRWPPQLAASREAPLSQRPHTQ 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 APPAPAAPPAVAPPA---GPRRAPPAVAAREPLPLPPQ-----PPAPPQQQPPPPQ 230

QY 65 GAGSPPE 71
   |||
Db 231 -PQPPE 236

RESULT 9
US-09-910-087-21
; Sequence 21, Application US/09910087
; Patent No. US20020055480A1
; GENERAL INFORMATION:
; APPLICANT: Koopman, Peter
;               Goodfellow, Peter
; TITLE OF INVENTION: SOX-9 GENE AND PROTEIN AND
;               USE IN THE REGENERATION OF BONE OR CARTILAGE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/910.087
; FILING DATE: 20-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM9714
; FILING DATE: 29-NOV-1994
; APPLICATION NUMBER: AU PM9835
; FILING DATE: 05-DEC-1994
; APPLICATION NUMBER: PCT/AU95/00799
; FILING DATE: 29-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
;

```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-010-087-21

Query Match      15.2%; Score 68.5; DB 10; Length 509;
Best Local Similarity 33.3%; Pred. No. 9.7;
Matches 24; Conservative 7; Mismatches 22; Indels 19; Gaps 6;

QY 11 GVPVAVHKHALHD---GLRVAGVGAHAHLPRW-----PPQLAASRRAPPPLSQS 59
Db 305 GVP-ATHGQVYTYGSGISSTAATPASAGHV--WMSKQAPPVPPQ----QPPQAPPAPQA 358
QY 60 PHTQAGSPPE 71
Db 359 PPQPQAA--PPQ 368

RESULT 10
US-09-823-240-9
; Sequence 9, Application US/09823240
; Patent No. US20020048813A1
; GENERAL INFORMATION:
; APPLICANT: Frank B. Gertler
; APPLICANT: James E. Bear
; APPLICANT: Jurgen Wehland
; APPLICANT: Joseph Loureiro
; TITLE OF INVENTION: Methods and Products for Regulating Cell
; FILE REFERENCE: Motility
; FILE REFERENCE: M0656/7064 (HCL)
; CURRENT APPLICATION NUMBER: US/09/823,240
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,564
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-823-240-9

Query Match      15.1%; Score 68; DB 10; Length 684;
Best Local Similarity 36.6%; Pred. No. 15;
Matches 15; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

QY 30 GPGAAAHLPRWPPPPQLAASRRAPPPLSQRPHRTQAGSP 70
Db 400 GPGGPPAPAPPPPPPSFGGAAGGPPPPAPPQMGAPP 440

RESULT 11
US-09-764-868-731
; Sequence 731, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 731
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-010-087-21

Query Match      14.9%; Score 67; DB 9; Length 721;
Best Local Similarity 34.7%; Pred. No. 20;
Matches 25; Conservative 4; Mismatches 29; Indels 14; Gaps 4;

QY 12 VPAVAVHKHALHDGLRQV-AGPGAAAHLPRWPPPPQLAASR-----REAPP-LSORP 60
Db 461 VPAKSRERLANGLHPVPMGPSGA--LPSDPAPCLPVKRGSPASPTSPDCPPALAPRP 517
QY 61 HRTQAGSPPE 72
Db 518 LSGQAGSPPE 529

RESULT 12
US-09-983-531A-4
; Sequence 4, Application US/09983531A
; Patent No. US20020147132A1
; GENERAL INFORMATION:
; APPLICANT: Fujisawa, Atsuko
; APPLICANT: Yamakawa, Toru
; APPLICANT: Shirakawa, Kamon
; APPLICANT: Chitose, Orii
; APPLICANT: Ogawa, Naoki
; TITLE OF INVENTION: Meltrins
; FILE REFERENCE: 11-22-99 sequence submission
; CURRENT APPLICATION NUMBER: US/09/983,531A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: JP 8-61756
; PRIOR FILING DATE: 1996-02-23
; PRIOR APPLICATION NUMBER: PCT/JP96/03017
; PRIOR FILING DATE: 1996-10-17
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Clone:
; OTHER INFORMATION: JM109(pBSmel-beta), mouse beta meltrin
; NAME/KEY: Unsure
; LOCATION: 574
; OTHER INFORMATION: Amino acid "Xaa" is unknown
US-09-983-531A-4

Query Match      14.9%; Score 67; DB 10; Length 920;
Best Local Similarity 36.6%; Pred. No. 27;
Matches 15; Conservative 5; Mismatches 19; Indels 2; Gaps 1;

QY 36 AHLPRWPPPPQLAASRRAPPPLSQRPHRTQAGSPPETNEKL 76
Db 789 SHPLRPPDYL--RVESPAPLSAHLNRAAGSSPEAGARI 827

RESULT 13
US-09-854-731-18
; Sequence 18, Application US/09854731
; Patent No. US20020120949A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/854,731
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/092,438
; PRIOR FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 625
```

; TYPE: PRT
; ORGANISM: Zea mays
US-09-854-731-18

Query Match 14.7%; Score 66.5; DB 10; Length 625;
Best Local Similarity 28.1%; Pred. No. 20;
Matches 27; Conservative 2; Mismatches 26; Indels 41; Gaps 5;
QY 5 AARPVAVHKHALHDLGVAGPGAAAHLPWRPPQLAAS-----49
DB 42 ATPRHKSGSTTPVHH-----QAATPGAAA-----WPSYPAGGASPLPAGVSPSPA 89
QY 50 -----RREAPPLSQRP-----RTQAGSPPE 71
DB 90 RSTPRFRFRPPPPSPAKHKATLAKRLGGGKPK 125

RESULT 14

US-09-858-664A-3
; Sequence 3, Application US/09858664A
; Patent No. US20020072491A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-3

Query Match 14.7%; Score 66.5; DB 10; Length 846;
Best Local Similarity 37.7%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 30 GPGAAAHLPWRPPQLAASRREA-PPLSQRPRTQAGSPPETNEKLTNPQV 81
DB 776 GFSQGAAPPSPKPHAAVFAVRSPPPGCAPEKRVPSAGGPPVLAEKARVPTV 828

RESULT 15

US-09-738-973-425
; Sequence 425, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 425

; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-425

Query Match 14.7%; Score 66.5; DB 10; Length 4019;
Best Local Similarity 31.1%; Pred. No. 1.7e+02;
Matches 19; Conservative 5; Mismatches 18; Indels 19; Gaps 2;
QY 31 PGAAAHLPWRPPQLAASRREAPPLSQRPRT-----OGAGSPPETNEKLTN 78
DB 2392 PMSAPAHLP LNPP-----RIQPIAQLPIKCTPACTVSNANPQSGPPRVEFDNN 2444
QY 79 P 79
DB 2445 P 2445

Search completed: November 26, 2002, 19:44:26
Job time : 46.2581 secs